

Review

Space, Time, and Fear: Survival Computations along Defensive Circuits

Dean Mobbs,^{1,2,*} Drew B. Headley,³ Weilun Ding,¹ and Peter Dayan^{4,5}

Naturalistic observations show that decisions to avoid or escape predators occur at different spatiotemporal scales and that they are supported by different computations and neural circuits. At their extremes, proximal threats are addressed by a limited repertoire of reflexive and myopic actions, reflecting reduced decision and state spaces and model-free (MF) architectures. Conversely, distal threats allow increased information processing supported by model-based (MB) operations, including affective prospection, replay, and planning. However, MF and MB computations are often intertwined, and under conditions of safety the foundations for future effective reactive execution can be laid through MB instruction of MF control. Together, these computations are associated with distinct population codes embedded within a distributed defensive circuitry whose goal is to determine and realize the best policy.

‘Everything is time-bound and space-bound’

Max Delbruck, 1949

Adaptive Computations Calculate the Predator’s Spatiotemporal Distance

Naturalists have offered compelling accounts of behaviors that are critical to survival [1,2] and include extensions of and support for Darwin’s selection theories, such as elegant mating displays, nurturing offspring, and foraging [3–6]. For prey, these behaviors occur in landscapes where lethal predators can emerge at any moment and they duly exhibit a rich spectrum of defensive behaviors suitable for potential and actual jeopardy [7]. Here, we consider how prey address the statistical and computational constraints associated with finding suitable actions to cope with dangerous environments in the face of substantial complexity and uncertainty. Predation is a particularly rich test case for the integration of hardwired and adaptive behavioral systems, and the psychological and neural structures involved provide clues about issues of human psychiatric importance, including aggression, **fear** (see [Glossary](#)), **anxiety**, panic, and rumination [8].

We divide the complexities of defensive behavior by considering the predatory environment, the knowledge an agent has about that environment, and then the policies it brings to bear in response. As first proposed in a model by Fanselow and Lester [7]; [Figure 1A](#), the spatiotemporal proximity or predatory imminence critically determines what actions might possibly be effective and how much time there is to choose among those actions. It thus constrains the neural circuitry involved [8–11]. In particular, the exigencies of threat place a critical burden on the speed and depth of decision-making. We review some neural responses to this challenge, including: hardwired policies; hardwired (Pavlovian) mappings from predicted affective evaluations of states to specific classes of action; restrictions on the optionality [12] of states (i.e., on the number of actions that are considered); off-line expansion of this optionality by forms of strategizing; and off-line consolidation of multistep planning into the equivalent of simple reactive reflexes ([Figure 1B](#)).

Highlights

Decisions to avoid or escape predators occur at different spatiotemporal scales, resulting in different computations and neural circuits.

At their extremes, surprising or proximal threats will reduce decision and state space and utilize model-free architectures, while distant threats allow increased information processing supported by model-based operations.

Model-free and model-based computations, however, are often intertwined. Furthermore, under conditions of safety the foundations for effective reactive execution in the future can be laid through model-based instruction of model-free control. Prospective planning can also be enabled.

Together, these computations reflect distinct population codes embedded within a distributed defensive circuitry whose goal is to determine and realize the best policy.

¹Department of Humanities and Social Sciences and Computation, California Institute of Technology, 1200 E. California Blvd, HSS 228–77, Pasadena, CA 91125, USA

²Neural Systems Program at the California Institute of Technology, 1200 E. California Blvd, HSS 228–77, Pasadena, CA 91125, USA

³Center for Molecular and Behavioral Neuroscience, Rutgers University – Newark, 197 University Avenue, Newark, NJ 07102, USA

⁴Max Planck Institute for Biological Cybernetics, 72076 Tübingen, Germany

⁵The University of Tübingen, Tübingen, Germany

*Correspondence: dmobbs@caltech.edu (D. Mobbs).



Threat Imminence Continuum and the Contiguity of Fear and Anxiety

Observations by naturalists suggest that defensive strategies [6] map onto a universal set of threat contexts, organized proximally to distally, which is crystallized in Fanselow and Lester's **threat imminence continuum** model (Figure 1A) [7, 13]. This canonical model of ecologically defined modes of threat, links defensive behaviors (e.g., flight, freeze) to levels of threat imminence. It provides a compelling platform from which to elucidate and decipher the role of the **defensive circuits** and computations in survival [6]. Further, it provides an objective multiscale dissection of fear, anxiety, and beyond, as well as linking old and new behavioral assays to different contexts of threat (Figure 2). This approach addresses potential issues with the 'black-and-white' fear-anxiety distinction, where the animal does not treat these as distinct either/or states but as waxing and waning along a gradient between the two extremes and that engage **population codes** distributed along a defensive circuit. Below we describe and extend the original threat imminence continuum model, but suggest contiguous subcategories of fear and anxiety defined by the spatiotemporal distance to danger [14, 15].

During the 'safe' or 'preferred activity' state, a prey might believe itself to be relatively safe, with no prospect of a predator. Its certainty of continuing to enjoy this context will depend on how volatile or changeable it thinks the world is. In the spirit of saving for a rainy day, many mammals will exploit this state to engage in preemptive defensive strategies such as positive niche construction, including the potentially reflexive or prespecified behavioral routines that create a safe environment (e.g., nests, burrows). A strategy observed in humans in this context is **intermittent anxiety**, which is a phasic form of anxiety and occurs when we are not preoccupied with the threat but leads us to occasionally think about the future threats we will or may encounter. Intermittent anxiety involves periodic checking for a state change and fleeting aversive cognition (e.g., flashes forward) about negative events that will, or could, happen in the distant future. If a distant threat is consciously imagined, humans can preempt danger through changing their environment or engaging in other protection strategies (e.g., staying in groups, carrying a weapon). It has yet to be shown whether animals show intermittent anxiety and, if so, whether this produce changes in protective behavior.

One step more dangerous is the pre-encounter context. Here, no threat is believed to be present, yet one may appear at any time. This expectation may be based on general priors about the environment or it may arise in locations near where the prey has specifically encountered a threat. The former inspires general avoidance, for instance of dark or open areas; the latter often results in place aversion. During foraging in potentially dangerous places, a defensive approach is often observed (e.g., in the way that potential food is approached) because of the potential risk of predatory attack [16]. In this case, **anticipatory anxiety** might be mandated rather than fear, along with vigilance, risk assessment, trepidation, and cautionary behaviors such as intermittent pauses, thigmotaxis, crouching, stretch [2], and eye darting [17]. Compared with the safe context, the pre-encounter threat context will also result in a reduction in foraging and mating behaviors. However, these behaviors will reappear over time through a process of extinction or better modeling of the environment (e.g., more precise, and less general, information about the potential dangers of the niche).

In the next context, called post-encounter threat, the prey has observed that a threat is present but believes itself to have not been detected or at least not to be the subject of untoward interest. Depending on how close the threat is (or indeed how close are other potential targets that would otherwise occupy the threat), this state can be associated with either anxiety or fear-like responses such as flight and freezing. We argue that a rather large range of behaviors is mandated in this case, depending on how much learning the prey has had the time or opportunity to do and/or consolidate and on the likely statistics of the conversion of the predator into the

Glossary

Affective prospection: the active, conscious use of episodic memory systems to imagine future threats or rewards.

Anticipatory anxiety: when there is no direct threat but a relatively high likelihood of encountering one in the immediate future (e.g., during pre-encounter threat).

Anxiety: a future-oriented emotional state associated with potential and uncertain threats.

Defensive circuit: a circuit that, given the immediacy and value of a threat, determines which policy/action to pick.

Defensive transition states: where agents switch from one defensive mode to another (e.g., from pre to post-encounter threat; see Figure 1A).

Dyna: an offline learning process in which, outside epochs of active behavior, replay or preplay samples are generated from a model of the world and are used to train an MF actor. Dyna allows a form of offline model inversion.

Encounter anxiety: when a threat is present but there is no interaction between the agent and the threat (e.g., during post-encounter threat).

Fear: an emotion that is associated with a present and identifiable threat. We use the term in conditions under which the agent is escaping a predator under circa-strike attack.

Hardwired responses: an innate choice of action or reflex that accompanies the learned prediction (for preparatory Pavlovian responses); a response that is not subject to instrumental learning based on its success or failure.

Intermittent anxiety: when there is no direct threat but there is a chance of encountering one in the distant future. Such phasic anxiety occurs during safety but is soon suppressed by cognitive control or reappraisal. For example, the knowledge that you will give a public lecture in 1 month will only cause anxiety when you think about it, but the lecture does not preoccupy your moment-to-moment thoughts.

Model-based (MB) methods: in MB planning, organisms exploit a (possibly learned) model of the environment to calculate prospectively the likely consequences of actions; for instance, by simulating possible future states. The model itself is typically learned and can be used for other operations (e.g., Dyna).

circa-strike context described below. Given less confidence in a benign outcome, post-encounter threat might generate **encounter anxiety**, which is time limited, involves strategizing avoidance involving anticipation of the threat's future behavior (will the predator attack me?), and avoidance (i.e., self-directed movement away from danger). This is separate from escape, which is driven by the predator's attack trajectories and environment contingences such as escape routes, proximity to a refuge, and protean escape. Post-encounter threat will also often be associated with an active form of freezing – a strategy that makes the prey harder to detect and gives it sufficient opportunity to gather relevant information and time to complete the required computations. This is compared with a passive form of quiescence that might also involve no movement, but also no substantial information gathering or processing. Along with other structures, these types of active and passive freezing may be controlled by the dorsal and ventral aspects of the periaqueductal gray (PAG), respectively (see [18]).

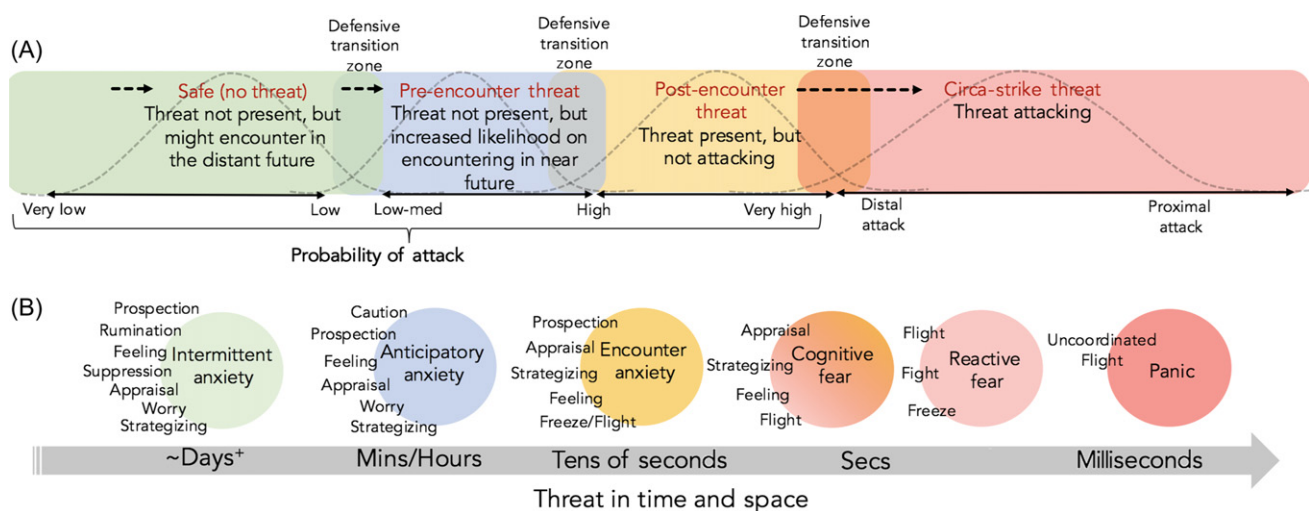
The most extreme danger is called circa-strike, which is when a predator is actively engaged in attacking the prey. In its original form, the circa-strike context is seen as a model for what inspires panic with consequences such as poorly thought-through choices. This may be because attempting to strategize in too sophisticated a fashion in such circumstances will slow the execution of more adaptive responses. Panic can certainly arise when the predator is overwhelmingly close and can be associated with hardwired (in the PAG) defensive behaviors such as defensive aggression and freezing. However, attacking threats will often appear at different distances, with quite different consequences. We thus consider circa-strike threat as the primary context that evokes different forms of fear and escape. First, if the attacking threat is distant, cognitive fear will be evoked as the animal has time to contemplate and strategize escape. In practice, this is closely related to (post)encounter anxiety, which we describe above. If the predator is somewhat closer but not yet within reach, the prey will experience what we call reactive fear and make **hard-wired** flight or fight responses aimed at inhibiting or fending off attacks from the predator [19,20]. This will also result in protean escape, which is driven by the trajectory of the predator's attack and often results in unpredictable flight (e.g., zigzagging, spinning; [21]). It is only when the predator is temporally (and thus spatially) close that a form of panic associated with uncoordinated and erratic escape might genuinely be the best response or at least better than no response.

Model-free (MF) methods: learn directly either what to do to maximize long-run return or learn value estimates of those long-run returns. MF methods acquire values by a bootstrapping process of enforcing consistency between successive estimates. MF methods acquire policies by stamping in 'what works' (i.e., what seems to lead to unexpected increases in values), creating something analogous to a reflex.

Population codes: neural codes studied at the population level.

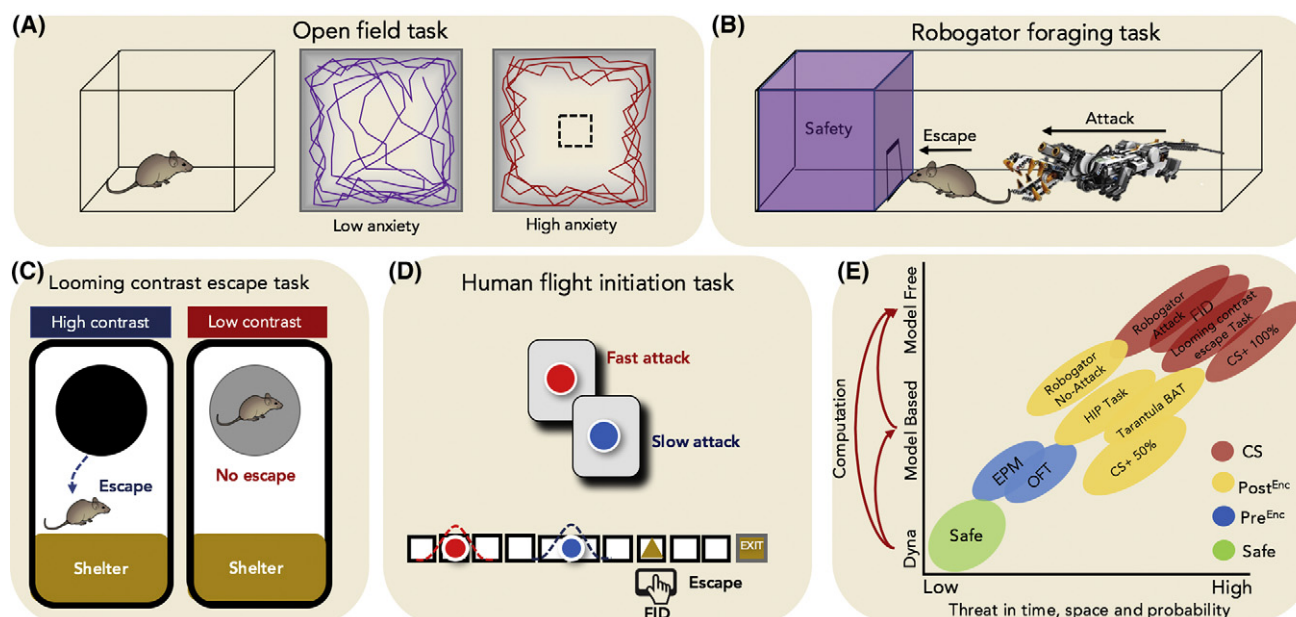
Threat imminence continuum: a canonical model of the ecologically defined levels of danger and how they relate to defensive behaviors (e.g., flight, freeze).

Time-uncertainty principle: as time into the future increases, so does the decision and state space and the variance in the predator's attack possibilities. This causes conflict and uncertainty in the defensive circuits.



Trends in Cognitive Sciences

Figure 1. Time, Space, Fear, and Anxiety. (A) The threat imminence continuum running from high- to low-imminence threats. (B) A hypothetical model of the different behaviors, cognitions, and emotions representing, panic, fear, and anxiety states that map onto the threat imminence continuum.



Trends In Cognitive Sciences

Figure 2. Examples of Behavioral Assays That Measure Different Spatiotemporal Scales of Danger. (A) The open field test (OFT) is used to assess pre-encounter threat, where there is potential to engage with a predator. In the OFT, rodents are typically placed in an unfamiliar box, where anxiety behaviors such as thigmotaxis are measured. (B) The predatory robot task, where the rat forages for pellets in the presence of a robogator. This task captures both the post-encounter and circa-strike levels of danger. For example, the rat will attempt to retrieve a pellet when the threat is not moving but this will switch over to circa-strike when the robogator looms toward the rat, causing the rat to escape back to the safety refuge. Behaviorally, rats will not collect pellets located beyond a certain distance but do retrieve pellets placed close to the refuge. (C) In the looming contrast escape task, the mouse reacts to innately aversive overhead expanding spots of different contrasts. High-contrast spots elicit escape, whereas low-contrast spots fail to cause escape. Reproduced, with permission, from [86]. Source of mouse stimuli: <https://scidraw.io/>. (D) Human flight initiation distance (FID) task, where subjects are told whether their decisions will result in high or low reward or shock. They are then presented with the image of the virtual predator where the color signals the attack distance (e.g., blue, fast; red, slow). After a short interval, the virtual predator appears at the end of the runway and slowly moves toward the subject's triangle. After an unspecified amount of time, the artificial predator will attack the subject's virtual triangle exit (i.e., attack distance). To escape, the subject must flee before the predator attacks. If the subject is caught, they will receive a tolerable, but aversive, shock to the back of the hand. Trials end when the predator reaches the subject or the exit. To motivate longer fleeing time, the task will include an economic manipulation, where subjects will obtain more money the longer they stay in the starting position and lose money the earlier they enter the safety exit. (E) Comparison of paradigms on two dimensions, (y) computation and (x) spatiotemporal distance. We present both conventional and new experimental assays. Abbreviations: EPM, elevated plus maze. Abbreviations: HIP, human intruder paradigm; CS+, conditioned stimulus; BAT, behavioral approach task. (See [88] for similar comparison between behavioral assays.)

The switches between cognitive and reactive fear and panic can happen in the range of milliseconds to seconds (Figure 1B).

Threat Imminence and Defensive Policies

The field of neural reinforcement learning (RL) provides a rich tapestry of formalizations of policies, which are systematic ways for the prey to act given its state in the world (Box 1). We next map this policy construction onto the imminence continuum, along with considering the psychological processes involved. In a preferred activity or 'safe' context, the appropriate policy for acting can be focused on appetitive instead of aversive outcomes. However, this is the perfect time for animals to engage in offline methods (Box 1) that burnish their aversively directed **model-free (MF)** and episodic policies, and even adjust their physical environment, to be able to cope better with threats when they arise. Animals may also have to perform computations to adjust their models of the world to accommodate or consolidate data on actual aversive interactions they had in more dangerous threat contexts but could not process online because of the danger. Choosing how much time and effort to spend in such offline operations may not be straightforward. Furthermore, there may also be negative consequences for an affective state of dwelling on highly negative, but most unlikely, outcomes, as in overexuberant rumination [22].

Box 1. Policy Construction

We consider formalizations of the nature and mechanisms of choice that come from the field of reinforcement learning (RL) [91]. These concern the prey's policy (also see [53,92]), which specifies which action to perform (or from what distribution this action is drawn) given a circumstance or state of the world, or at least beliefs about the world. This choice of action should have a beneficial effect on the expected longevity of the prey starting from the state; we call this longevity the long-run action value of executing that action at the state. States themselves also have values, averaging over the policy that is applied. Various methods are known for determining and constructing policies, some of which operate mainly online as the animal faces its current circumstance; others use information collected in the past, either directly from the world or indirectly from simulating or imagining the world. Different methods are appropriate for different degrees of threat imminence. We therefore first outline the possibilities and then indicate how they might be used in practice.

Two important dichotomies in RL associated with issues of learning and estimation are MB versus MF methods of planning and control [93] and hardwired or Pavlovian versus instrumental choice. MB (or goal-directed) computations typically reflect prospective simulation of the tree of possible future actions and outcomes (e.g., with the threat). These computations are typically slow but can accurately reflect all of the information the animal has. Furthermore, to the extent that learning of any sort is feasible (prey that are not born knowing at least something about defending themselves may not prosper long), acquiring a model is typically straightforward. MF (or habitual) choice is simpler and faster [30], reflecting instead learning from previous actual or simulated experience. However, the MF choice requires substantial learning to be accurate, because it works by bootstrapping something that can be particularly challenging given the heterogeneity of situations involving survival and the life-threatening cost of collecting information [94]. Fortunately, even only partially accurate MF values can be helpful. For instance, various algorithms replace distal parts of the prospective simulation tree that would normally be used to evaluate future states in a MB manner with MF estimates of intermediate values. The shallower the ultimate MB component of the tree, the faster the choice can be; the deeper into the tree that MF replacement occurs, the less the potential damage of any inaccuracy. The more pressing the need to make a decision, the shallower the use of MF values (e.g., [95]). There are also various other ways of blending MB and MF control (e.g., [95]).

Pavlovian control specifies generic actions (e.g., freezing) in the light of the valence of actual or MB or MF predicted outcomes (predator) and sometimes particular actions associated with particular outcomes. In both cases, prespecification obviates the problem of choice, but at the expense of inflexibility. By contrast, instrumental control allows the MB or MF specification of arbitrary actions in the light of their past or predicted future contingent effects. Pavlovian and instrumental control compete and cooperate.

Since MF control is faster than MB, there is particular advantage in training it offline to make it accurate. Computational methods such as Dyna suggest that what amounts to a model of the environment can internally generate, as rumination or replay, simulated observations that recapitulate what the animal might actually experience [91,96]. These can train the MF system so that it realizes a policy itself that is accurate as well as being computationally simple [91,96]. To put this another way, a MF policy can be seen as a form of inverse of the model, with offline simulation being used as a form of consolidation or compilation arranging model inversion, preparing the prey to be a more effective defender in any time of need. Sample generation, which can also be seen as a form of scenario testing, can credibly operate only offline, when MB and MF systems are not engaged in controlling behavior. In animals, relevant epochs are periods of quiet wakefulness and may be represented by default-mode neural activity [63,97]. Note, however, that although preplay and replay are widely observed under these conditions, and indeed some forms of them are known to be important for control [98], whether and how they operate in the precise case of defense is as yet unclear. It also remains to be seen to what extent humans enjoy greater abilities than rodents and other animals to taste virtual failure or success without risking deadly consequences of testing it out in the real world.

We also note that a key role is played by optionality – that only relatively few options might sometimes even be considered [12]. It is typically advantageous to have more options at one's disposal, but only if choice remains fast (**time-uncertainty principle**). MB strategizing and counterfactual 'prospective' reasoning (both of which are prominent in rumination) are ways of judiciously managing the set of potential options. Doing this offline, perhaps in consort with a form of Dyna, would allow the benefits of boosted or suppressed optionality to be provided at only limited cost to the speed of planning and control when a threat is actually present.

Along with MB and MF control, there is also evidence for a more episodic-like control strategy [57,99] which was originally thought of as reusing singular experiences from the world to determine appropriate actions in the face of very limited experience. However, one can imagine adding to the buffer of possible episodic suggestions by engaging offline with a form of **affective prospection** of future threat encounters (i.e., conscious imagination of future encounters with a threat) as a form of Dyna for episodic control. This would allow the agent to play through threat scenarios and predict the best action to select (i.e., if the predator does A, I will do B) or to engage in longer-term defensive policies such as niche construction.

In pre-encounter threat, there will still be time to engage in online **model-based (MB)** planning and risk assessment, taking advantage of the extra accuracy of this method of evaluation to determine things such as safe routes for foraging that avoid known or potential places where

predators appear. Furthermore, it may be possible to perform judicious probing actions, collecting more information to reduce uncertainty about the current state of the environment to determine the likely nature and appearance of any danger. Compared with safe contexts, computations (including preplay) are likely to be focused on the current state of the animal in its environment rather than potential future states.

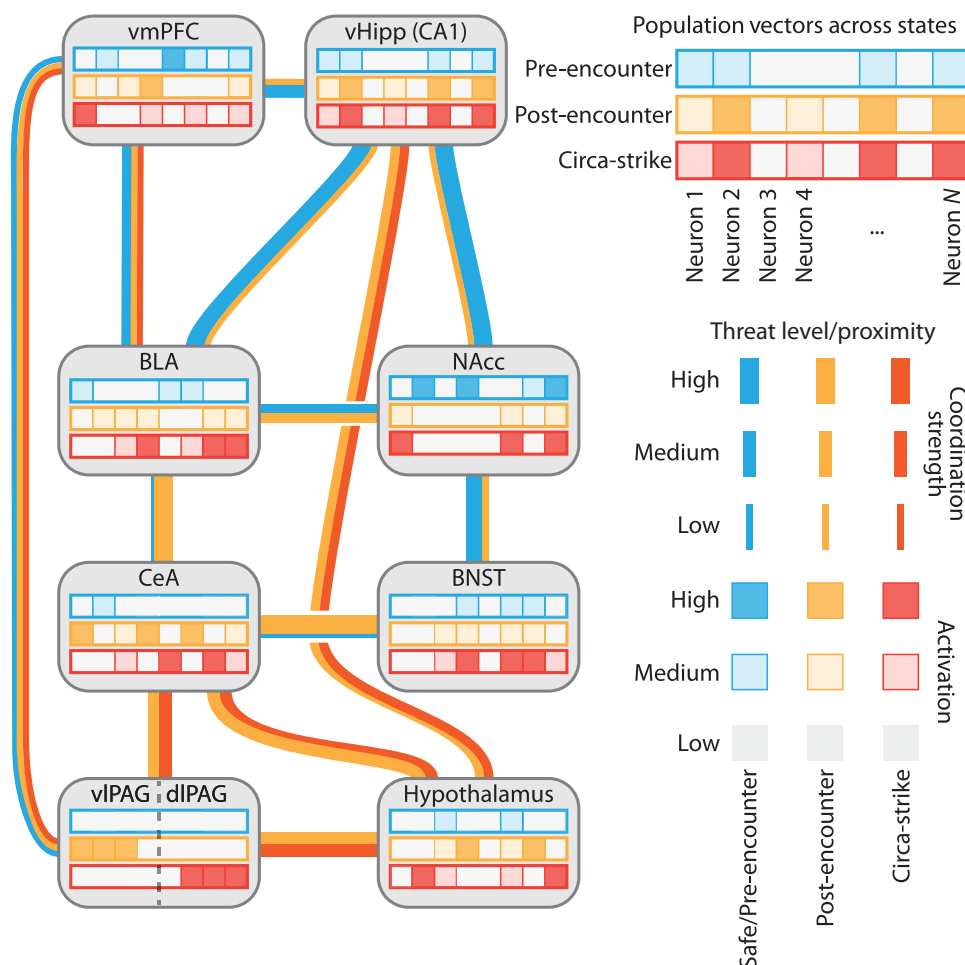
During post-encounter threat, the time horizon for the prey will normally have substantially shortened. This favors MF values (or at least MF contributions to MB values). We might expect hardwired, Pavlovian responses to observations of, and predictions about, the threat, notably including freezing, approach to safe refuges, and perhaps forms of urgent assessment that quantify the threat's magnitude and nature. However, more environmentally adaptive MF instrumental responses that can take advantage of either online learning of previous successful approaches or offline strategizing and model inversion will be of potentially even greater value. Any instrumental action can become MF through experience, either in the world, or via **Dyna**-style mental simulation. This could include the choice of a particular place of safety in an environment or the adoption of a particular sort of aggressive display that efficiently deters a predator of a given type. This is why it is beneficial for animals to have engaged in making this system accurate during safe contexts. If the post-encounter threat is sufficiently distant, then, mirroring pre-encounter threat, there may be time for MB planning (or hybrid combinations of MB and MF planning) to assess the threat's motive and strategy and to choose to gather necessary information and/or construct an appropriate escape plan.

Finally, during circa-strike threat, predators typically use stealth, camouflage, or speed making attack unexpected or proximal, leaving little time to think or prepare. Prey then need to have at their disposal a set of quick and deterministic defensive reactions. These can also be seen as (hopefully effective) hardwired Pavlovian heuristics consistent with a specialized neural structure that has privileged access to low-optionality control. The startle reaction, a stereotyped jump to a rapid threat like a loud burst of sound, is one of the simplest but most effective forms of defense and involves a hardwired brainstem circuit. More complex defensive reactions that reflect external risk assessment (e.g., external stimuli causing sensory motor responses rather than internal stimuli that reflect memory and cognition) take the form of flight, fight, and freezing reaction patterns and are observed across taxa. Even rather basic responses can depend on relatively sophisticated information; for instance, fleeing requires the maintenance and use of knowledge about where in space safety lies [23].

An important upside of Pavlovian control in post-encounter threat and circa-strike contexts is its speed; however, this partly arises from its important downside of inflexibility and thus potential inefficiency. One of the most central roles for cognitive control is to suppress a headlong rush to immediate, and potentially counterproductive, defensive actions. It also offers a flexible capacity to appraise danger in the context of relevant sources of information [23,24].

Defensive Circuits, Polices, and Population Coding

The questions that remain concern the implementation of these various defensive policies and policy-generating algorithms in the rich panoply of neural circuits that are involved in defense (Figure 3). As is typical for neural processing, we can describe the computations that occur in terms of the way that information (e.g., about threats) is represented and then the ways that these representations change in the light of new external data (e.g., sensory input) or internal calculations (e.g., determining the value of some particular action). Representation in the brain almost ubiquitously involves large and distributed populations of neurons; we therefore start by considering the structurally different sorts of population codes that arise in the case of defense



Trends in Cognitive Sciences

Figure 3. Population Codes and Threat Imminence. A simplified example of the population codes are represented in the defensive circuits discussed here. Schematized population vectors highlight the evolution the ensemble activity across the threat proximity axis. For instance, hippocampus shows relatively stable coding of the context as threat increases, amygdala shows separate populations for safety versus circa-strike, and the central amygdala is particularly active during post-encounter and circa-strike periods. Hypothalamus refers to hypothalamic nuclei that have been implicated in defensive behaviors, such as the lateral hypothalamus (e.g., [89]) and the ventromedial hypothalamic nucleus (e.g., [90]). Abbreviations: BLA, basolateral amygdala; BNST, bed nucleus of the stria terminalis; CeA, central nucleus of the amygdala; dIPAG, dorsolateral periaqueductal gray; NAcc, nucleus accumbens; vHIPP, ventral hippocampus; vPAG, ventrolateral periaqueductal gray; vmPFC, ventromedial prefrontal cortex.

[25] before turning to the intricate, and densely coupled, circuits that realize and operate over these codes. MB and MF methods of planning, and indeed Pavlovian and instrumental methods of control, can exploit whatever representations are available to support reliable (and learnable) correlates of long-run values. However, the circuits involved in these different methods partly overlap and can exhibit different forms of population codes [26], as we outline below.

The population codes used across defensive circuits can be loosely grouped into three types. First, sensory processing areas in the brainstem and sensory thalamic nuclei systematically tile the space of peripheral receptors (e.g., somatotopy, tonotopy; see Figure 1 in Box 2). In addition, behavior-related hypothalamic and brainstem nuclei (such as the PAG) encode stereotyped

behavioral repertoires or specific actions [27]. Second are structures that group sets of stimuli or behaviors into categories irrespective of peripheral details. These include the basolateral amygdala, which encodes the valence of stimuli and behaviors [25,28] and the entorhinal cortex, which captures aspects of the geometry of the environment [29]. Last are areas with highly conjunctive representations of the categories found in the second group. Both the hippocampus and the prefrontal cortex (PFC) stand out in this regard. The population code in the hippocampus is hierarchically organized to encode context, position, objects, and their values, among other things [30]. The PFC [for this section we are primarily focused on the prelimbic/infralimbic cortex in rodents and the ventromedial PFC (vmPFC) in humans] contains conjunctive coding as well, but without an apparent hierarchical organization [31,32]. However, one should keep in mind that hierarchical relationships do exist in the recruitment of PFC subregions [33,34] especially in humans during social processing [35]. This diversity of responses and lack of an obvious organization, known as mixed selectivity, is thought to support the processing of complex situations and novel rules [32].

It is worth considering what these different population codes say about the circuitry that gives rise to them. Simple properties of inputs such as a loud noise or a change in illumination, which are transparently represented in the activity of peripheral sensory neurons, will be able to drive behaviors reliably via even the earliest sensory processing regions in subcortical and cortical structures. By contrast, although single neurons can perform sophisticated computations [36,37], the rich conjunctions of sensory features that define ethologically relevant stimulus categories can generally be extracted only by multiple layers of population representations and re-representations. These are realized in feedforward and recurrent neural circuits, which are known to be able to realize efficient [38] and universal function approximators. Areas providing categorical representations that are invariant to peripheral details benefit from this depth of processing and are situated higher in the hierarchy.

The representations in both hippocampus and PFC can be distinct from the immediate sensory environment or its direct implications for values or actions. Hippocampal CA1 and CA3 exhibit both spontaneous and task-related replay of spatial trajectories [39] during active and passive

Box 2. Defensive Codes as Population Codes

The behavioral response to a threat is often specific to the situation, reliable, and time locked. By contrast, single-neuron activity is usually noisy, inconsistent, and spontaneously ongoing. This divide can be partly bridged by considering the collective activity of many neurons as the functional unit that supports behavior. At its most simple, if a threatening sound occurs, a downstream neuron can pool together the activity of many unreliable upstream neurons to determine when that stimulus is present. Crucially, with a population code, the upstream neurons do not need to respond the same way to all stimuli. Indeed, they may respond differently to all other potential stimuli. However, their suitably weighted aggregate activity can still reliably signal the threatening stimulus. When multiple stimuli similarly drive these neurons, then a downstream neuron that pools their activity will not be able to tell those stimuli apart. Sometimes this will correspond to semantically meaningful categories, like all cues that signal an impending shock. These correlations between neurons in their responsiveness will determine what is easiest to decode (i.e., situations that drive neurons together will be conveyed with greater efficacy than those that are only sparsely expressed).

Consider how these codes would appear to an experimentalist recording from single neurons during the presentation of a tone that had previously been paired with a shock. Each of the structures in Figure 1 contains some neurons that responded during the shock-associated cue, but the functional interpretation of this response by itself is unclear. To understand what a particular neuron is signaling, one must record its response across a variety of conditions (see bottom row). For instance, a sensory-driven neuron in the cochlear nucleus responds to the stimulus irrespective of whether it signals an aversive outcome. Likewise, a behavior-driven neuron in the PAG will respond whenever a particular behavior is emitted and be insensitive to the cue that elicits it. At the level of categorical representations, a neuron in the basolateral amygdala may respond to all aversive cues. For PFC, with mixed selectivity, the response could be just for a particular cue when it signals a threat in a specific context. Sampling the response in only one condition gives a distorted picture of its meaning.

Even if we consider the variety of circumstances that drives a single neuron in a particular region, it is important to remember that it is the collective activity of many neurons that is responsible for activating a downstream target. The frequency of an acoustic cue is well represented in the population activity of the cochlear nucleus, while its aversiveness is not. As mentioned above, features that are shared across many neurons will be easier to decode by downstream neurons. Thus, it is the pattern of population codes across a variety of circumstances that gives clear insights into the function of a region.

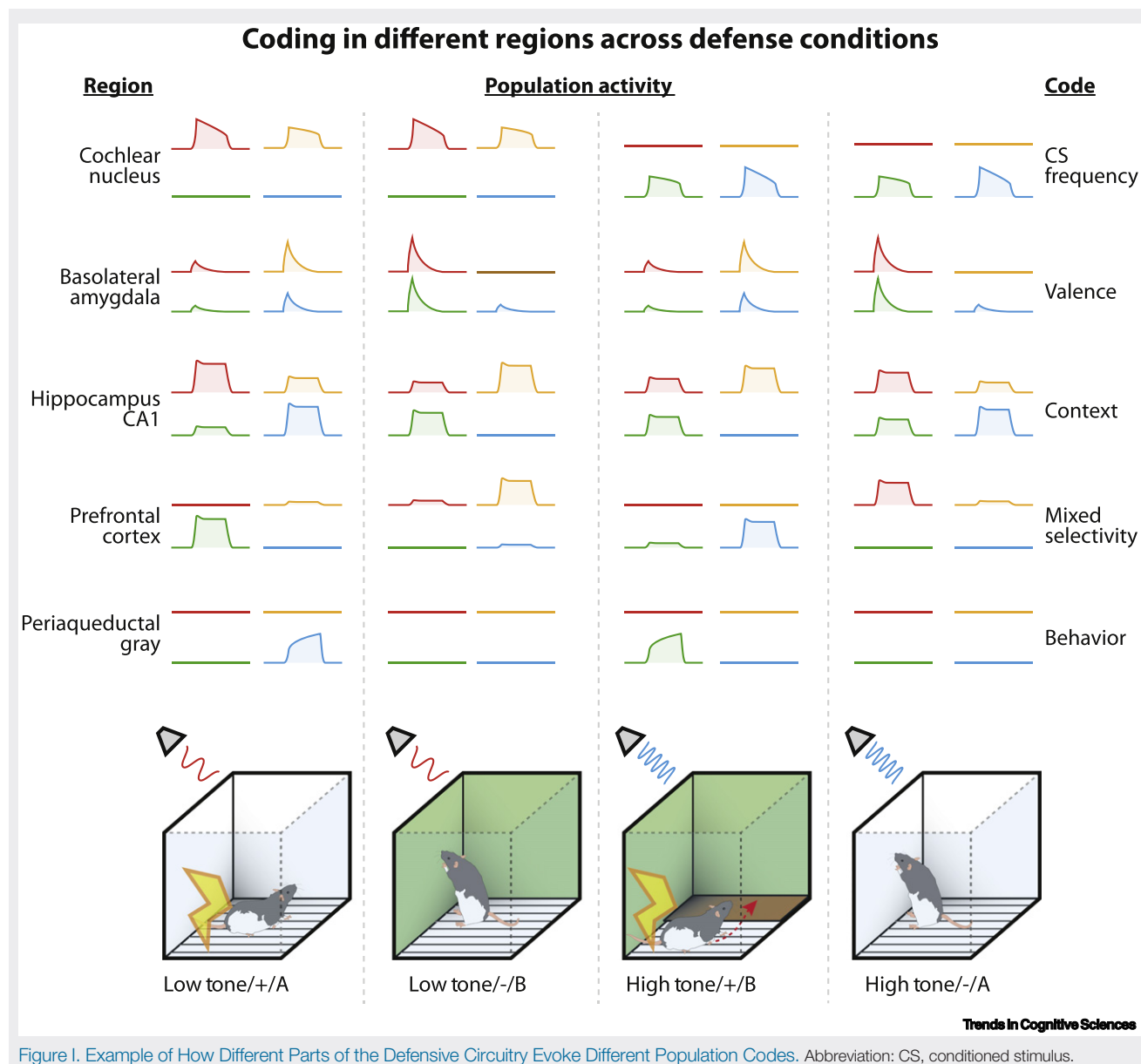


Figure 1. Example of How Different Parts of the Defensive Circuitry Evoke Different Population Codes. Abbreviation: CS, conditioned stimulus.

behavioral states. In the PFC, neurons can persist in their response to a stimulus even after it has disappeared [40]. Such representations can underpin forms of working memory, allowing information that is no longer present in the environment to affect behavior. Such representations depend on local recurrent excitatory connectivity (e.g., [41]).

By contrast, low-level sensory and behavioral effectors are particularly conducive to those Pavlovian policies that entail linking simple stimulus conditions with fixed actions. One example is acoustic startle, which can be mediated entirely via a brainstem pathway where auditory signals arising from cochlear nuclei activate the nucleus reticularis pontis caudalis, which in turn projects to motor neurons in the spinal cord [42]. Since loudness is already encoded at the level of the cochlea,

there is little need for refined stimulus representations or categorization. Likewise, since the behavior is highly stereotyped, it can be generated by effector-level motor circuits. Other regions have a role that appears to relate to affective aspects of aversive stimulus processing. In particular, the insular cortex is part of a parallel pathway for conveying pain information to the amygdala [43] that in humans is associated with subjective awareness of aversive stimuli [44].

When acoustic startle is potentiated by acquired threatening cues, brainstem circuits are no longer sufficient. Instead, areas encoding the learned threat levels for cues or contexts are important, such as the amygdala and the bed nucleus of the stria terminalis (BNST) [45,46], which have been suggested as substrates representing MF values. Yet more complex situations that demand flexibility in the response (e.g., freeze or avoid), or where the threat levels of stimuli depend on time and context, recruit the hippocampus [46] and PFC [47] for prospective MB calculations. What determines these shifts in populations is the spatial and temporal distance to the threat (Figure 3).

However, it is important to note that population activities can be exploited in diverse ways by different downstream processes. First, within a region, functionally distinct ensembles may project to different targets. This appears to be the case in the hippocampus, with neurons representing anxiety projecting to the lateral hypothalamic area while those supplying contextual information project to the basal amygdala [48,49]. Alternatively, individual neurons could represent multiple types of information, but the codes arising from their population activity would occupy separate subspaces. Neurons in motor cortex respond to both movement preparation and production, but the population codes for those two states are orthogonal to one another [50] and thus could be separated by a downstream neuron [51]. Last, downstream areas may be sensitive to particular conjunctions of activities spanning multiple regions. In that case, the effect a particular ensemble has is conditioned on the activity in other regions. These three possibilities are not mutually exclusive and thus they may be used in combination so that the same regions can support different defensive behaviors across the threat imminence continuum.

Circuits Involved in Pre-encounter Threat

The substrates of MB control in general include regions of the PFC, the hippocampus, and the amygdala [52]. There are human and rodent data suggesting that distal threat results in increased BOLD signal and neuronal firing in the vmPFC. This has led some to suggest that the vmPFC is involved in deliberative strategizing [6,53] and distinguishes between safe and aversive locations [54]. Recent evidence suggests that the vmPFC contains 'strategy selective cells' that are involved in eliciting and inhibiting movement such as freezing or fleeing [55], putatively via interaction with the amygdala, ventral hippocampus, and PAG [54]. There is also evidence that the basolateral amygdala (BLA) and ventral hippocampus provide memory and prospection input by evoking past experiences to envision what the threat will do in the future [56]; episodic memory may be of particular importance because of the paucity of information and experience in many cases of threat [57]. This circuitry has been linked to anxiety [58–60], which is characterized by future anticipation of threat. This is evident during hippocampal sharp-wave ripple events in awake rats, where trajectories leading to a shock zone are replayed prior to its avoidance (putative risk assessment [61]) and during sleep, where ripples coordinate BLA and hippocampal ensembles activated by a replayed aversive trajectory [62]. In humans, state representation may provide a similar function by representing the specific information needed to maximize avoidance (see [63]). Further, the conscious feeling of fear might also emerge from the use of working memory for prospective planning [64].

Circuits Involved in Post-encounter Threat

As we noted above, a mixture of MB and MF mechanisms are likely to be involved in situations where a threat is present. For example, many animals will exhibit pre-encounter behavioral

patterns when a threat arises, yet will switch from MB planning to MF calculations as it nears. With this transition, defensive behaviors in the post-encounter situation may become increasingly dependent on the BLA [65], BNST [66], hypothalamus [67], and PAG [68] and less so on the PFC and hippocampus. The BLA is capable of driving different post-encounter defense-related behaviors through its efferent targets, such as freezing via PFC (prelimbic subregion [69]) and central amygdala [70], passive avoidance via the central amygdala [71], and active avoidance via the nucleus accumbens [72]. The projection-defined subsets of BLA neurons are selected by extrinsic excitatory afferents [73] and inhibit one another through local inhibitory connectivity [74]. One possibility, therefore, is that the selection of these pathways arises from relatively simple local circuit mechanisms, which accords with a tilt toward MF processing. However, the PFC may contribute as well. During periods where the presence of threats is ambiguous, such as entering a new context, PFC is able to directly affect the evocation of defensive behaviors via its descending projections to the PAG [75]. Activation of this projection can also increase pain thresholds, which may be crucial for preparing to engage in defensive behavior [76].

Various structures have been implicated in the implementation and acquisition of MF policies in the case of defense. Tovote and colleagues [68] showed the existence of an inhibitory pathway that extended from the central nucleus of the amygdala (CeA) to the ventrolateral PAG. This results in the CeA producing freezing by ultimate disinhibition of ventrolateral PAG excitatory outputs to pre-motor targets (e.g., the magnocellular nucleus of the medulla). The CeA is also particularly interesting in the rapid selection of flight or freeze responses. For example, Fadok and colleagues [77] show that CeA cells expressing corticotropin-releasing factor (CRF+) mediate flight while activation of somatostatin-positive (SOM+) neurons results in freezing, and local inhibitory connections between CRF+ and SOM+ neurons mediated the rapid selection of flight or freezing. Thus, this circuit may be evoked at either distal or proximal post-encounter threat and circa-strike attack.

The switch from MB to MF may also be reflected at the level of the PFC and hippocampus. These regions can drive the same defensive behaviors through different pathways. As fear memory consolidates, the PFC switches from activating freezing through a BLA to central nucleus pathway and toward a midline thalamus to central nucleus route [78]. Concordantly, consolidation also changes the ensembles in PFC that are activated during fear-memory retrieval [79]. Like the PFC, the hippocampus has multiple routes to the same defensive behavior. Ventral CA1 neurons projecting to the basal amygdala support contextual freezing, while those projecting to the central nucleus contribute to freezing during fear renewal [49]. Avoidance behaviors during anxiogenic situations, such as the elevated plus maze, depend on ventral CA1 projections to the lateral hypothalamus [48].

Together, this suggests that the vmPFC and hippocampus remain involved in the coordination of defensive behaviors throughout the threat imminence continuum. However, the ensembles within them change if the threat comes too close. MF and hardwired circuits (e.g., amygdala) play a role in more reflexive reactions. The post-encounter mode is interesting from an experimental perspective because it provides a way to study the arbitration between deliberative MB and reflexive MF computations.

Defensive Circuit Activity during Circa-Strike Attack

The ubiquity of defensive reactions such as flight and freezing supports their adaptive success; furthermore, the neural circuits involved in these behaviors are similar in a variety of species. They are linked to a set of higher midbrain tegmentum regions such as the PAG. This is the core neural substrate involved in defensive behaviors and different behaviors map onto functionally distinct neuronal columns of the structure: the dorsomedial, dorsolateral, lateral, and ventrolateral PAG [80]. For instance, rodent studies suggest that regions of the rostral

dorsomedial/dorsolateral PAG (dm/dlPAG) [80] are involved in flight during circa-strike, while stimulating the ventrolateral PAG leads to behavioral quiescence in the form of freezing, which is typically a post-encounter response. The arbitration between these divergent behaviors arises from the differential activation of afferent pathways, such as the central nucleus for freezing [68] and the superior colliculus for escape [23], along with inhibitory circuits between PAG nuclei [81]. Given the close correspondence between specific PAG pathways and behaviors, their activation by particular ethologically relevant stimuli (e.g., [23]), and the fact that the ensembles in the PAG that encode these behaviors rarely overlap [82], it stands as an example of a region implementing MF processing.

More recent optogenetics-based experiments have shown that activation of PAG neurons can induce defensive responses including running, freezing, and avoidance [68,83]; by contrast, lateral/ventral lateral inhibition impairs PAG post-encounter defensive responses [68]. These suggest that PAG is not only an output circuit but also a center implicated in the control of threat responses. Moreover, the connectivity between the PAG and other regions also mediates how an agent reacts to threat. The ventrolateral and dorsal PAG have been shown to convey unconditional stimulus information to the BLA, which directs both innate and learned fear responses [84]; the dorsal PAG interacts with higher cortical centers, such as the PFC, to mediate both pre-encounter and post-encounter defensive behaviors [55,85,86]. Further, recent evidence suggests that the superior colliculus is key in detecting the saliency of the threat, while the dorsal PAG's glutamatergic neurons encode the escape choice [23]. Humans seem to possess similar circuits; for instance, as a tarantula is placed closer to a subject's foot in the MRI scanner, the PAG monotonically ramps up its activity [10] and more recent work has implicated the PAG in fast, but not slow, escape decisions [20].

Concluding Remarks

We have proposed that adaptive defensive decisions are directly linked to the spatiotemporal distance of the threat (see [Outstanding Questions](#)). The defensive circuitry beneficially allows an agent to adapt its behavior flexibly based on the level of threat imminence, yet a reflexive MF system supports appropriate choice in the face of the often-unpredictable nature of predation. Offline Dyna-like preparation, in which MB systems teach MF systems appropriate courses of action, is used to prepare the ground for making adaptive decisions later when there is little time to think; that is, since, under proximal threat, deep information processing (MB, rumination) is inefficient, adaptive defensive decisions will be able to rely on hardwired Pavlovian and MF systems. Our proposed framework provides further insights into various different ways that these circuits and computations can go awry, where having too high priors for the occurrence of predators or unreasonably stringent requirements for the probability of being able to get to safety [22,87] may underlie psychiatric disease.

Acknowledgments

This work was supported by US National Institute of Mental Health grant 2P50MH094258 and a Chen Institute Award (P2026052) (to D.M.) and the Max Planck Society and the Humboldt Foundation (to P.D.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the authors' funders.

References

1. Lima, S.L. and Dill, L.M. (1990) Behavioral decisions made under the risk of predation: a review and prospectus. *Can. J. Zool.* 68, 619–640
2. Cooper Jr., W.E., Blumstein, D.T., eds (2015) *Escaping from Predators: An Integrative View of Escape Decisions*, Cambridge University Press
3. Mobbs, D. et al. (2018) Foraging for foundations in decision neuroscience: insights from ethology. *Nat. Rev. Neurosci.* 19, 419–427
4. Darwin, C. (1979) *The Origin of Species: Complete and Fully Illustrated*, Gramercy Books
5. LeDoux, J. (2012) Rethinking the emotional brain. *Neuron* 73, 653–676
6. Mobbs, D. et al. (2015) The ecology of human fear: survival optimization and the nervous system. *Front. Neurosci.* 9, 55
7. Fanselow, M.S. and Lester, L.S. (1988) A functional behavioristic approach to aversively motivated behavior: predatory

Outstanding Questions

What is the adaptive function of conscious feelings of fear and anxiety [64,100]? The functional basis of these feelings, which LeDoux and Hoffman have proposed are 'the essence of emotion' in humans, may play a role in the volitional control of movement [101,102]. This may be particularly relevant to altering behavior based on more distant, prospective goals. A question is how do such conscious feelings influence MB computations?

Are the cortical regions involved in conscious defensive states separate from survival circuits or part of a contiguous defensive circuit? To answer this question, we will need to create new paradigms and techniques that use time and space to investigate populations of neurons across the entire brain.

What are the regions of the brain that orchestrate the switch between defensive states and how do these differ across species? We expect this control circuitry to consume estimates of the spatiotemporal distance of the danger.

Are cognitive fear and encounter, anticipatory, and intermittent anxiety part of the same circuitry?

How do defensive circuits, as described here, go wrong in affective psychopathology? Are some affective disorders related to inept estimation of spatiotemporal dangers?

- imminence as a determinant of the topography of defensive behavior. In *Evolution and Learning*, pp. 185–211, Erlbaum
8. Perusini, J.N. and Fanselow, M.S. (2015) Neurobehavioral perspectives on the distinction between fear and anxiety. *Learn. Mem.* 22, 417–425
 9. Mobbs, D. et al. (2007) When fear is near: threat imminence elicits prefrontal-periaqueductal grey shifts in humans. *Science* 317, 1079–1083
 10. Mobbs, D. et al. (2010) Neural activity associated with monitoring the oscillating threat value of a tarantula. *Proc. Natl. Acad. Sci. U. S. A.* 107, 20582–20566
 11. Mobbs, D. et al. (2009) From threat to fear: the neural organization of defensive fear systems in humans. *J. Neurosci.* 29, 12236–12243
 12. Simsek, O. et al. (2016) Why most decisions are easy in Tetris – and perhaps in other sequential decision problems, as well. *Proc. Mach. Learn. Res.* 48, 1757–1765
 13. Blanchard, R.J. and Blanchard, D.C. (1989) Antipredator defensive behaviors in a visible burrow system. *J. Comp. Psychol.* 103, 70–82
 14. Shackman, A.J. and Fox, A.S. (2016) Contributions of the central extended amygdala to fear and anxiety. *J. Neurosci.* 36, 8050–8063
 15. Fox, A.S. et al. (2015) Extending the amygdala in theories of threat processing. *Trends Neurosci.* 38, 319–329
 16. McNaughton, N. and Corr, P.J. (2004) A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neurosci. Biobehav. Rev.* 28, 285–305
 17. Blanchard, D.C. (2017) Translating dynamic defense patterns from rodents to people. *Neurosci. Biobehav. Rev.* 76, 22–28
 18. Brandão, M.L. et al. (2008) Different patterns of freezing behavior organized in the periaqueductal gray of rats: association with different types of anxiety. *Behav. Brain Res.* 188, 1–13
 19. Mobbs, D. (2018) The ethological deconstruction of fear(s). *Curr. Opin. Behav. Sci.* 24, 32–37
 20. Qi, S. et al. (2018) How cognitive and reactive fear circuits optimize escape decisions in humans. *Proc. Natl. Acad. Sci. U. S. A.* 115, 3186–3191
 21. Driver, P.M. and Humphries, D.A. (1970) Protean displays as inducers of conflict. *Nature* 226, 968–969
 22. 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
 23. Evans, D.A. et al. (2018) A synaptic threshold mechanism for computing escape decisions. *Nature* 558, 590–594
 24. Evans, D.A. et al. (2019) Cognitive control of escape behaviour. *Trends Cogn. Sci.* 23, 334–348
 25. Headley, D.B. et al. (2019) Embracing complexity in defensive networks. *Neuron* 103, 189–201
 26. Morrison, S.E. et al. (2011) Different time courses for learning-related changes in amygdala and orbitofrontal cortex. *Neuron* 71, 1127–1140
 27. Franklin, T.B. (2019) Recent advancements surrounding the role of the periaqueductal gray in predators and prey. *Front. Behav. Neurosci.* 13, 60
 28. Kyriazi, P. et al. (2018) Multi-dimensional coding by basolateral amygdala neurons. *Neuron* 99, 1315–1328.e5
 29. Solstad, T. et al. (2008) Representation of geometric borders in the entorhinal cortex. *Science* 322, 1865–1868
 30. McKenzie, S. et al. (2014) Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron* 83, 202–215
 31. Rigotti, M. et al. (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497, 585–590
 32. Rigotti, M. et al. (2010) Internal representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. *Front. Comput. Neurosci.* 4, 24
 33. Badre, D. and Nee, D.E. (2018) Frontal cortex and the hierarchical control of behavior. *Trends Cogn. Sci.* 22, 170–188
 34. Wutz, A. et al. (2018) Different levels of category abstraction by different dynamics in different prefrontal areas. *Neuron* 97, 716–726.e8
 35. Chambon, V. et al. (2017) Neural coding of prior expectations in hierarchical intention inference. *Sci. Rep.* 7, 1278
 36. Koch, C. (2005) *Biophysics of Computation: Information Processing in Single Neurons*, Oxford University Press
 37. Branco, T. and Häusser, M. (2011) Synaptic integration gradients in single cortical pyramidal cell dendrites. *Neuron* 69, 885–892
 38. Eldan, R. and Shamir, O. (2016) The power of depth for feedforward neural networks. In *JMLR Workshop Conf. Proc.* (49), pp. 1–34
 39. Carr, M.F. et al. (2011) Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nat. Neurosci.* 14, 147–153
 40. Fuster, J.M. (1973) Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. *J. Neurophysiol.* 36, 61–78
 41. Mante, V. et al. (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503, 78–84
 42. Yeomans, J.S. and Frankland, P.W. (1995) The acoustic startle reflex: neurons and connections. *Brain Res. Brain Res. Rev.* 21, 301–314
 43. Shi, C. and Davis, M. (1999) Pain pathways involved in fear conditioning measured with fear-potentiated startle: lesion studies. *J. Neurosci.* 19, 420–430
 44. Critchley, H.D. et al. (2002) Volitional control of autonomic arousal: a functional magnetic resonance study. *Neuroimage* 16, 909–919
 45. Davis, M. and Shi, C. (1999) The extended amygdala: are the central nucleus of the amygdala and the bed nucleus of the stria terminalis differentially involved in fear versus anxiety? *Ann. N. Y. Acad. Sci.* 877, 281–291
 46. Chowdhury, N. et al. (2005) Dorsal hippocampus involvement in trace fear conditioning with long, but not short, trace intervals in mice. *Behav. Neurosci.* 119, 1396–1402
 47. Gilmartin, M.R. et al. (2014) Prefrontal cortical regulation of fear learning. *Trends Neurosci.* 37, 455–464
 48. Jimenez, J.C. et al. (2018) Anxiety cells in a hippocampal–hypothalamic circuit. *Neuron* 97, 670–683.e6
 49. Xu, C. et al. (2016) Distinct hippocampal pathways mediate dissociable roles of context in memory retrieval. *Cell* 167, 961–972.e16
 50. Kaufman, M.T. et al. (2014) Cortical activity in the null space: permitting preparation without movement. *Nat. Neurosci.* 17, 440–448
 51. Poirazi, P. et al. (2003) Pyramidal neuron as two-layer neural network. *Neuron* 37, 989–999
 52. Coutureau, E. and Killcross, S. (2003) Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behav. Brain Res.* 146, 167–174
 53. Bach, D.R. and Dayan, P. (2017) Algorithms for survival: a comparative perspective on emotions. *Nat. Rev. Neurosci.* 18, 311–319
 54. Adhikari, A. et al. (2011) Single units in the medial prefrontal cortex with anxiety-related firing patterns are preferentially influenced by ventral hippocampal activity. *Neuron* 71, 898–910
 55. Halladay, L.R. and Blair, H.T. (2015) Distinct ensembles of medial prefrontal cortex neurons are activated by threatening stimuli that elicit excitation vs. inhibition of movement. *J. Neurophysiol.* 114, 793–807
 56. Bulley, A. et al. (2017) Thinking about threats: memory and prospection in human threat management. *Conscious. Cogn.* 49, 53–69
 57. Lengyel, M. et al. (2005) Matching storage and recall: hippocampal spike timing-dependent plasticity and phase response curves. *Nat. Neurosci.* 8, 1677–1683
 58. Adhikari, A. (2014) Distributed circuits underlying anxiety. *Front. Behav. Neurosci.* 8, 112
 59. Fung, B.J. et al. (2019) Slow escape decisions are swayed by trait anxiety. *Nat. Hum. Behav.* 3, 702–708
 60. Adhikari, A. et al. (2010) Synchronized activity between the ventral hippocampus and the medial prefrontal cortex during anxiety. *Neuron* 65, 257–269
 61. Blanchard, D.C. et al. (2011) Risk assessment as an evolved threat detection and analysis process. *Neurosci. Biobehav. Rev.* 35, 991–998

62. Girardeau, G. *et al.* (2017) Reactivations of emotional memory in the hippocampus–amygdala system during sleep. *Nat. Neurosci.* 20, 1634–1642
63. Schuck, N.W. and Niv, Y. (2019) Sequential replay of nonspatial task states in the human hippocampus. *Science* 364, eaaw5181
64. LeDoux, J.E. and Brown, R. (2017) A higher-order theory of emotional consciousness. *Proc. Natl. Acad. Sci. U. S. A.* 114, E2016–E2025
65. Moscarello, J.M. and Maren, S. (2018) Flexibility in the face of fear: hippocampal–prefrontal regulation of fear and avoidance. *Curr. Opin. Behav. Sci.* 19, 44–49
66. Goode, T.D. and Maren, S. (2017) Role of the bed nucleus of the stria terminalis in aversive learning and memory. *Learn. Mem.* 24, 480–491
67. Kunwar, P.S. *et al.* (2015) Ventromedial hypothalamic neurons control a defensive emotion state. *Elife* 4, e06633
68. Tovote, P. *et al.* (2016) Midbrain circuits for defensive behaviour. *Nature* 534, 206–212
69. Senn, V. *et al.* (2014) Long-range connectivity defines behavioural specificity of amygdala neurons. *Neuron* 81, 428–437
70. Jimenez, S.A. and Maren, S. (2009) Nuclear disconnection within the amygdala reveals a direct pathway to fear. *Learn. Mem.* 16, 766–768
71. Tye, K.M. *et al.* (2011) Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature* 471, 358–362
72. Ramirez, F. *et al.* (2015) Active avoidance requires a serial basal amygdala to nucleus accumbens shell circuit. *J. Neurosci.* 35, 3470–3477
73. McGarry, L.M. and Carter, A.G. (2017) Prefrontal cortex drives distinct projection neurons in the basolateral amygdala. *Cell Rep.* 21, 1426–1433
74. Krabbe, S. *et al.* (2018) Amygdala inhibitory circuits regulate associative fear conditioning. *Biol. Psychiatry* 83, 800–809
75. Rozeske, R.R. *et al.* (2018) Prefrontal–periaqueductal gray-projecting neurons mediate context fear discrimination. *Neuron* 97, 898–910.e6
76. Huang, J. *et al.* (2019) A neuronal circuit for activating descending modulation of neuropathic pain. *Nat. Neurosci.* 22, 1659–1668
77. Fadok, J.P. *et al.* (2017) A competitive inhibitory circuit for selection of active and passive fear responses. *Nature* 542, 96–100
78. Do-Monte, F.H. *et al.* (2015) A temporal shift in the circuits mediating retrieval of fear memory. *Nature* 519, 460–463
79. DeNardo, L.A. *et al.* (2019) Temporal evolution of cortical ensembles promoting remote memory retrieval. *Nat. Neurosci.* 22, 460–469
80. Bandler, R. *et al.* (2000) Central circuits mediating patterned autonomic activity during active vs. passive emotional coping. *Brain Res. Bull.* 53, 95–104
81. Tovote, P. *et al.* (2015) Neuronal circuits for fear and anxiety. *Nat. Rev. Neurosci.* 16, 317–331
82. Deng, H. *et al.* (2016) Periaqueductal gray neuronal activities underlie different aspects of defensive behaviors. *J. Neurosci.* 36, 7580–7588
83. Assareh, N. *et al.* (2016) The organization of defensive behavior elicited by optogenetic excitation of rat lateral or ventrolateral periaqueductal gray. *Behav. Neurosci.* 130, 406–414
84. Kim, E.J. *et al.* (2013) Dorsal periaqueductal gray–amygdala pathway conveys both innate and learned fear responses in rats. *Proc. Natl. Acad. Sci. U. S. A.* 110, 14795–14800
85. Vander Weele, C.M. *et al.* (2018) Dopamine enhances signal-to-noise ratio in cortical–brainstem encoding of aversive stimuli. *Nature* 563, 397–401
86. Lefler, Y. *et al.* (2020) The role of the periaqueductal gray in escape behavior. *Curr. Opin. Neurobiol.* 60, 115–121
87. Huys, Q.J.M. *et al.* (2015) Depression: a decision-theoretic analysis. *Annu. Rev. Neurosci.* 38, 1–23
88. Shackman, A.J. *et al.* (2016) Dispositional negativity: an integrative psychological and neurobiological perspective. *Psychol. Bull.* 142, 1275–1314
89. Marcinkiewicz, C.A. *et al.* (2016) Serotonin engages an anxiety and fear-promoting circuit in the extended amygdala. *Nature* 537, 97–101
90. Miller, S.M. *et al.* (2019) Divergent medial amygdala projections regulate approach–avoidance conflict behavior. *Nat. Neurosci.* 22, 565–575
91. Sutton, R.S. and Barto, A. (1998) *Reinforcement Learning: An Introduction*, MIT Press
92. Bishop, S.J. and Gagne, C. (2018) Anxiety, depression, and decision making: a computational perspective. *Annu. Rev. Neurosci.* 41, 371–388
93. Doya, K. (1999) What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw.* 12, 961–974
94. Daw, N.D. (2018) Are we of two minds? *Nat. Neurosci.* 21, 1497–1499
95. Keramati, M. *et al.* (2016) Adaptive integration of habits into depth-limited planning defines a habitual-goal-directed spectrum. *Proc. Natl. Acad. Sci. U. S. A.* 113, 12868–12873
96. Mattar, M.G. and Daw, N.D. (2018) Prioritized memory access explains planning and hippocampal replay. *Nat. Neurosci.* 21, 1609–1617
97. Greicius, M.D. *et al.* (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* 100, 253–258
98. Jadhav, S.P. *et al.* (2012) Awake hippocampal sharp-wave ripples support spatial memory. *Science* 336, 1454–1458
99. Bornstein, A.M. *et al.* (2017) Reminders of past choices bias decisions for reward in humans. *Nat. Commun.* 8, 15958
100. LeDoux, J.E. and Pine, D.S. (2016) Using neuroscience to help understand fear and anxiety: a two-system framework. *Am. J. Psychiatry* 173, 1083–1093
101. Fanselow, M.S. (2018) The role of learning in threat immittance and defensive behaviors. *Curr. Opin. Behav. Sci.* 24, 44–49
102. LeDoux, J.E. and Hofmann, S.G. (2018) The subjective experience of emotion: a fearful view. *Curr. Opin. Behav. Sci.* 19, 67–72