

1 This pre-print is now published open-access in
2 Neuroscience and Biobehavioral Reviews:

3
4 <https://doi.org/10.1016/j.neubiorev.2025.106213>

Critical intelligence: computing defensive behaviour

Jules Brochard¹, Peter Dayan^{2,3}, Dominik R Bach^{1,4}

¹University of Bonn, Transdisciplinary Research Area Life and Health, Center for Artificial Intelligence and Neuroscience, Bonn, Germany

²Max Planck Institute for Biological Cybernetics, Tübingen, Germany

³University of Tübingen, Tübingen, Germany

⁴Wellcome Centre for Human Neuroimaging, University College London, UK

Abstract: 182 words

Main text: 6500 words

References: 203

Figures: 4

Text boxes: 1

Acknowledgements

DRB receives funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. ERC-2018 CoG-816564 ActionContraThreat) and through the iBehave Network, which is sponsored by the Ministry of Culture and Science of the State of North Rhine-Westphalia, Germany. The Hertz Chair for Artificial Intelligence and Neuroscience in the Transdisciplinary Research Area Life and Health, University of Bonn, is funded as part of the Excellence Strategy of the German federal and state governments. The Wellcome Centre for Human Neuroimaging is supported by core funding from the Wellcome (203147/Z/16/Z). PD is funded by the Max Planck Society and the Humboldt Foundation, and is a member of the Machine Learning Cluster of Excellence, EXC number 2064/1 – Project number 39072764 and of the Else Kröner Medical Scientist Kolleg "ClinbrAI: Artificial Intelligence for Clinical Brain Research". The authors thank Paolo Domenici for commenting on an initial version of the manuscript.

34 Competing Interests

35 The authors declare no competing interests.

Abstract

Characterising the mechanisms underlying naturalistic defensive behavior remains a significant challenge. While substantial progress has been made in unravelling the neural basis of tightly constrained behaviors, a critical gap persists in our comprehension of the circuits that implement algorithms capable of generating the diverse defensive responses observed outside experimental restrictions. Recent advancements in neuroscience technology now allow for an unprecedented examination of naturalistic behaviour. To help provide a theoretical grounding for this nascent experimental programme, we summarise the main computational and statistical challenges of defensive decision making, encapsulated in the concept of critical intelligence. Next, drawing from an extensive literature in biology, machine learning, and decision theory, we explore a range of candidate solutions to these challenges. While the proposed solutions offer insights into potential adaptive strategies, they also present inherent trade-offs and limitations in their applicability across different biological contexts. Ultimately, we propose series of experiments designed to differentiate between these candidate solutions, providing a roadmap for future investigations into the fundamental defensive algorithms utilized by biological agents and their neural implementation. Thus, our work aims to provide a roadmap towards broader understanding of how complex defensive behaviors are orchestrated in the brain, with implications for both neuroscience research and the development of more sophisticated artificial intelligence systems.

Introduction

To survive and thrive, biological agents must avoid harm from immediate threats. Successful defence requires selecting appropriate actions—often very rapidly— using the correct motor commands to execute them, adapting to dynamic situations, and learning to avoid danger in the future. The high stakes, time pressure, and an unrelenting evolutionary arms race, render this challenge significantly different from those encountered in other domains of dynamic behaviour (Cisek and Pastor-Bernier, 2014; Gordon et al., 2021), such as foraging and hunting, mating and parental care, or social and cultural activities. Autonomous artificial agents face similar problems when seeking to ensure survival (e.g. in airplanes) or dangerous events such as collisions (e.g. in self-driving cars) (Egner, 2009; Khatib, 1987; Todorov, 2004), associated with the substantial field of real-time systems engineering (Kopetz and Steiner, 2022). In this review, we adopt the term "critical intelligence" to encapsulate the challenge of navigating this unique dilemma. In political and military contexts, where "intelligence" typically refers to knowledge, the term "critical intelligence" denotes information that demands

immediate attention and action to avert danger. In contrast, within the fields of neuroscience and machine learning, "intelligence" is understood as a capability. Thus, we define critical intelligence as the ability to make decisions that are simultaneously complex, swift, and accurate.

Understanding how the brain achieves critical intelligence would address a fundamental neurobiological question, could help advance our understanding of maladaptive avoidance in clinical conditions, and inform the design of safe autonomous agents. However, experimental research into this question has long been hampered by practical constraints. For example, classical neuroscience paradigms make only a carefully selected set of behaviours meaningful, by restricting the environment and/or the available motor actions. This reduction in complexity simplifies the problem the agent faces (Fanselow, 1994; Fanselow and Lester, 1988; Mobbs et al., 2009) to the extent that it can be solved by combining a small number of decision strategies. In particular, it will admit largely pre-specified, reflexive or model-free algorithms (Bach and Dayan, 2017; LeDoux and Daw, 2018; Mobbs et al., 2020). In contrast, outside these constraints, a much larger array of defensive behaviors is reported in field research and behavioral studies (Domenici et al., 2011a, 2011b; Evans et al., 2019). This has highlighted the sophistication and complexity of critical intelligence in many species (Domenici et al., 2011a, 2011b; Evans et al., 2019) including humans (Sporrer et al., 2023), using naturalistic virtual reality (Brookes et al., 2023).

Novel neuroscience methods such as neural imaging in freely moving animals (Klioutchnikov et al., 2023) and humans (O'Neill et al., 2025; Roberts et al., 2019; Snider et al., 2013; Topalovic et al., 2020) now allow relaxing constraints and investigating the neural generation of critical intelligence in more naturalistic situations. In this review, we seek to provide a conceptual scaffold to guide this nascent research field. This scaffold rests on two pillars. The first pillar is an analysis of the statistical and computational challenges that any biological or artificial agent would encounter when responding to a defensive situation. The second pillar is a collection of candidate solutions for these challenges—many of which have not been investigated for defensive scenarios in biological agents. Often, several computational solutions exist for the same problem, necessitating experiments to disambiguate these possibilities—we outline some such experiments in textbox 1.

Important structure for our discussion follows the levels of analysis programme laid out by Marr & Poggio (1977), whilst emphasising that the best, i.e., the optimal computational solution, can be the enemy of the good. i.e., whatever a neurally-realizable algorithm can produce in real time (figure 1). In other words, the joint demands of rapidity and sufficiency determine not only the type of behaviour required – i.e., the computational solution – but also the algorithm(s) and implementation(s) by which it can be found. This duality runs through our review – and implies that constraints of the (neural)

substrate can significantly impact the ethological computations that evolution could have crafted and honed. Nevertheless, the levels of analysis provide a useful organizing principle, not the least because it allows us to discuss some of the multiplicity of solutions rather transparently.

After summarising the background of our approach, we first discuss the challenges associated with finding optimal defensive strategies (i.e., the computational level). These strategies might be stochastic and are often couched in structured, symbolic, terms. Next, we discuss challenges related to planning actual defensive behaviour (i.e., the algorithmic level), drawing on descriptions from neuroscience and psychology. Finally, we discuss learning without direct feedback, which poses its own computational demands.

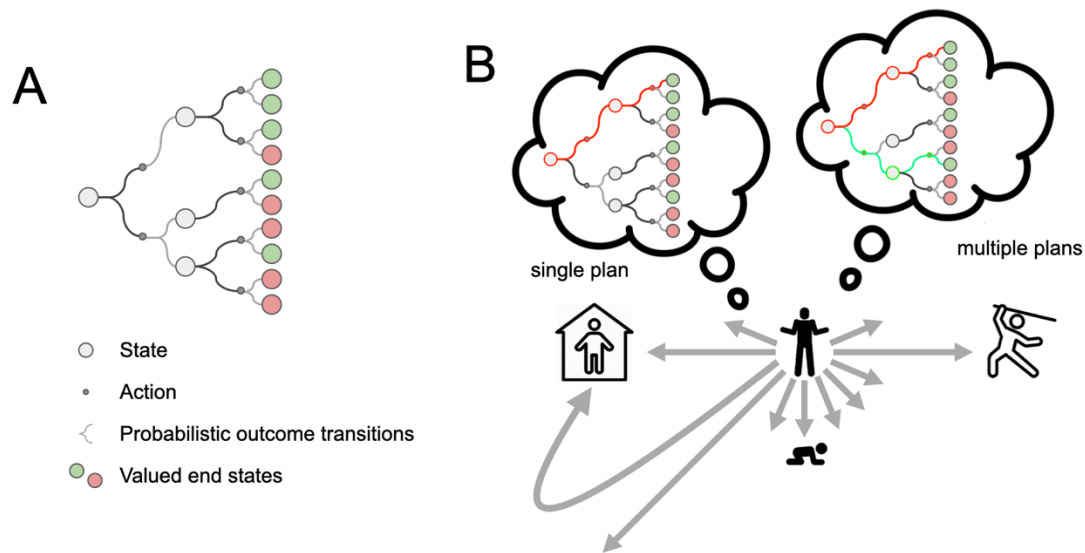


Figure 1: Critical intelligence. A. Decision trees are a classical way of depicting a multi-step decision problem. The agent's state (circles) is influenced by its previous actions (black lines and small dots), but actions can have non-deterministic outcomes (grey lines). The time horizon is indicated by the horizontal extent of the decision tree. B. An agent, confronted by a threat, needs to decide quickly between multiple action options, many of which require delicate execution planning to be successful (longer arrows), while others can be executed rapidly. Some of these actions might be appropriate only under the current circumstances whilst others might be more robust when the environment changes. With infinite time, an optimal action could be found, but at least some of the options require rapid initiation. What the agent actually does will therefore depend not only on what is considered the most appropriate action (traditionally a computational level question), but also on which action can be found within the limited time available. This, in turn, depends on which algorithm the agent is actually using, which for many common scenarios and many species is as yet unknown. For example, an agent might plan just one action for the current situation (red trace, left-hand side) but could also

make multiple action plans simultaneously, which might take more time but could potentially account for incoming information or changes of the environment, and thus be useful for the future (green trace, right-hand side).

Box 1: Disambiguating algorithms for critical intelligence using behavioral perturbations experiments

To understand the neural algorithms that compute behaviour, we suggest distilling candidate mechanisms in behavioural experiments. Such experiments should approximate the complexity of the problem space that the agent uses in a natural environment – they should be naturalistic. To distinguish candidate mechanisms that result in similar behaviour, experiments should perturb the natural environment in ways that make some algorithms fail, but not others.

Objective: Do agents maximise the probability of survival over a certain time horizon, or expected capture time? For example, in defending against a faster opponent in open territory, will they trade a close and insecure shelter against a far and even less secure shelter?

Time horizon: Does initial energy expenditure change, when the duration of a flight path is extended beyond the anaerobic time horizon?

Adaptability: Are agents able to change initial responses when an opponent abandons its pursuit, or when an additional opponent interferes?

Closed-loop control: Are agents able to integrate new information into their behaviour, and if so, which type of information? For instance, what happens if its field of view abruptly widens or clears, or if the likely results of its ongoing response change suddenly.

Robust actions: In scenarios with multiple escape options, how early do agents commit to one of them?

Pre-planning: When a preferred escape route closes at the same time that a threat appears, does the agent still approach it?

Hierarchical planning: Does the selection of escape paths depend only on shelter location or also on detailed characteristics, for example any time-consuming motor actions required to access a shelter?

Anytime planning: In a prolonged predator-prey stand off, is the response to a sudden lurch of the predator as fast at different time points?

Factorisation: When different response options have very different opportunity windows, can agents integrate this to find the overall best response?

Learning from past experience: To what extent does peer play, observation, or near-catch experience generalise to threat encounters with varying features such as speed, path, angle of approach, parts of the environment, or goals.

Background

Biological agents are constrained by their neural hardware, which has evolved by happenstance in response to the demands of specific ecological niches. Ample anatomical and functional evidence suggests that many species are equipped with dedicated neural circuits that generate specific defensive behaviours (Bach and Dayan, 2017). Examples are giant Mauthner cells eliciting an ultra-fast escape reflex in many fish and amphibians (Bullock, 1984), the oligosynaptic startle circuitry producing a protective withdrawal reflex in many mammals (Davis, 2006), and a specialised threat learning system in mammals (LeDoux, 1995). All of these systems deal (almost) exclusively with situations of threat. While such examples are abundant across the animal kingdom, they generally produce narrow and often short-term responses. It remains unclear how the continuous and complex defensive behaviour that characterizes many species in their natural habitat (Domenici et al., 2011b, 2011a; Evans et al., 2019; Sporrer et al., 2023) is generated. For the example of threat learning, there is ongoing debate about the extent to which this process is additionally supported by a generic reward learning system outside the most constrained laboratory paradigms (Laing et al., 2024). When it comes to more sophisticated and apparently goal-directed defensive behaviors, and their well-known adaptation over multiple time scales (Lima and Dill, 1990), the underlying neural systems remain largely elusive.

From a purely theoretical standpoint, many solutions for generating such complex behaviours appear possible. It seems reasonable to expect, on the one hand, that dedicated neural circuits subserve at least a subset of defensive behaviours, possibly beyond the systems already known. On the other hand, for organisms that occupy diverse habitats such as humans, it would be implausible to expect a dedicated neural circuit for every defensive situation they may encounter, implying that some functions will be realized by more general-purpose systems. Examples of such circuits also exist in model organisms, such as the tectospinal system in fish, which controls goal-directed locomotion not only during escape but across different behavioural goals (Saitoh et al., 2007). Interestingly, this raises a possibility that different behaviours are controlled by different systems, but might co-occur in some situations. Indeed, research in humans has suggested that what appears like a coherent escape sequence is subserved by different underlying systems (Sporrer et al., 2023). Such systems can

typically be differentiated only by directed experimental manipulations, not by mere observation alone. These manipulations require theoretical considerations, including hypotheses about the inner workings of candidate algorithms or neural circuits that control behavior.

In the remainder of this article, we seek to provide a roadmap for this work by a thorough analysis of the unique statistical and computational challenges that agents face in defensive situations, with the goal of inspiring theoretical and empirical research on how different species solve these challenges. We also draw potential solutions from a rich computer science, engineering, and computational neuroscience literature. These should be regarded as selected set of candidates, to be confirmed or refuted in empirical research. We are not yet in a position to propose implementations of these solutions in neural circuits, largely because the novelty of the required experimental neuroscience methods has limited the extent of research on the neural substrate. Thus, instead, we focus on algorithms. Of course, we do not claim that these algorithms are “implemented” in neural circuits in a directly representational sense (Brette, 2019). Instead, the required computations might well be carried out in distributed dynamical systems with no notion whatsoever of the quantities that appear in an abstract algorithm (Churchland and Shenoy, 2024). Thus, the algorithms we propose here provide an idealised, abstract description of how behaviour is generated, and a benchmark against which to compare the output of a (simulated) neural model (Bialek, 2022). They can also be used for comparison to artificial agents in a signature testing approach (Taylor et al., 2022).

Two of the critical terms and concepts that appear throughout the review are that of a **state** and of a **policy**. The state characterizes the situation of the agent in terms of the environment, its history of its own interactions with that environment, and internal factors such as hunger and thirst. The policy is a mapping from states to actions, and, as the key determinant of success and failure of the agent, is the ultimate target of information processing. Critical tradeoffs exist between the speed, efficacy and complexity of computing different policies. As a result, both artificial and natural agents enjoy several structurally different mechanisms for generating, and then arbitrating between, policies.

The simplest policy is to map states of the (perceived) **external** environment directly to actions – this is often called a **reflexive algorithm**. As an especially simple case, a **reflex** (e.g. the startle reflex) is a policy that takes only a small part of the environment into account and is largely insensitive to the remaining sensory input (Yeomans et al., 2002). An agent might switch between different reflexive algorithms depending on some internal variables – this type of control has been described as “behavioural algorithms” (Hein et al., 2020). In general, reflexive algorithms might be hard-wired but they could also be learned from experience. One way of learning them is model-free reinforcement learning. This simply increases the probability of an action in a certain environmental state if this action leads to a desirable outcome (Thorndike, 1911). Reflexive algorithms can also be acquired by

‘fixing’ the output of initially more sophisticated learning mechanisms (Gershman, 2020). These types of transferred reflexive policy are often called "habits" (Dickinson and Balleine, 1994; Dolan and Dayan, 2013).

In contrast, a **reflective algorithm** entails an explicit consideration of the agent's goals, the outcomes of its actions, and the utility of these outcomes in relation to the agents' goals (Mattar and Lengyel, 2022). The emergence of reflective algorithms entails learning and updating a model of the environment, often termed model-based learning.

There are three important considerations for this picture. First, the distinction between reflexive and reflective algorithms should be thought of as ends of a spectrum rather than a pair of categories (Collins and Cockburn, 2020). For example, even "simple" reflexes such as the startle reflex appear modulated by internal estimates of the state such as prior expectations or opportunity costs (Bach, 2015a). Second, much work in the field assumes that states and actions are discrete, but the actual environment is continuous. Generalising discrete state algorithms to continuous environments is the subject of intense theoretical research (e.g., Smith et al., 2020). Finally, the common assumption of a completely known state is a simplification, and a fuller characterization is provided by partially observable Markov decision processes (POMDP). This is an important construct in computer science, albeit one without fast or general solutions.

The computational level: Defining optimal defensive response strategies

Even with ample resources, identifying the best, or even a good, defensive strategy poses crucial challenges inherent in the problem setting. The computational level provides a third-person perspective, where we are not concerned with how the solutions are actually achieved (figure 2). The main challenges are what the objective of a defensive action should be and when it should be emitted, how to behave under uncertainty, and how to outwit one's opponent.

Specifying the objective and optimising response timing

The first challenge is **combining multiple objectives** (figure 2a). The primary goal of the agent is survival, but agents will often have to balance short-term and long-term survival. For example, in foraging under risk of predation, organisms may trade a small chance of short-term death for resources that facilitate long-term survival (Ydenberg and Dill, 1986). Thus, the challenge here is balancing contradictory objectives.

Different computational solutions have been proposed (e.g., Enkhtaivan et al., 2023). A popular solution is scalarisation – mapping the different goals onto a scalar objective function, such as a loss function in Bayesian decision theory (Berger, 2013), a utility function in expected utility theory (Bernoulli, 1738; Savage, 1954; von Neumann and Morgenstern, 1944), a drive function in homeostatic reinforcement learning (Keramati and Gutkin, 2011, 2014), or an assumed objective in allostatic control (Sennesh et al., 2022). However, as yet, there exists no normative way of constructing such objective functions from the organism’s needs. Instead, they are usually inferred from behaviour (e.g. choice preferences) (e.g., von Neumann and Morgenstern, 1944) or derived from physiological data (e.g. capacity) (e.g., Sennesh et al., 2022) data. Besides scalarisation, a second solution is selective attention – addressing different goals one at a time (Enkhtaivan et al., 2023). Finally, heuristic objective functions have been proposed, such as capture time (Weintraub et al., 2020), probability of success (Wynn et al., 2015), or minimum threat distance (Cooper Jr., 1997; Kawabata et al., 2023; Soto et al., 2015).

A related consideration is the time horizon over which the objective is optimized. For example, extended defensive responses pose particular demands on **energy expenditure**. To avoid interception, an agent may have to perform intense movements rapidly. However, repeated or prolonged efforts will lead to exhaustion and later vulnerability (Guinet et al., 2007). With limited energy reserves, agents must balance them carefully to ensure the quality of their performance and secure their overall escape (Mandralis et al., 2021). Behaviourally, at least for simple actions, it appears that energy demands of a movement are integrated with the rewards it affords (Cos et al., 2011; Shadmehr et al., 2016), thus making it likely that the agent has a notion of energy expenditure in their decision-making.

Computationally, the time horizon is implicitly encoded in the discount factor of (homeostatic) reinforcement learning models; but in this framework it is usually treated as a free model parameter and inferred from behaviour, rather than specifying how the agent generates this parameter. There is work on optimizing the energy demands of single actions on the fly and in a distributed system (Todorov, 2004). This concept typically does not afford *a priori* estimates of the energy cost of different complete defensive responses. On the other hand, an agent might attempt to calculate the energy cost of chunks of sub-sequences (Balleine and Dezfouli, 2019; Tomov et al., 2020), to allow overall energy needs to be approximated.

The next challenge is determining optimal **response timing** (figure 2b). For example, when circled by a bird of prey, an agent might quickly decide on a dodge response, but may choose to withhold it until the bird enters the ballistic attack phase during which it cannot change its own trajectory any longer (Evans et al., 2019, see for a similar situation Bach, 2015b). In contrast, responding early might be beneficial in other situations but energetically costly (Dayan, 2012; Nord et al., 2017; Shadmehr, 2020).

Moreover, as often the case in embodied decision-making, different actions might enjoy different windows of opportunity, such that their optimal timing differs (Cisek and Pastor-Bernier, 2014; Gordon et al., 2021).

Computationally, an agent needs a form of temporal prediction to integrate its beliefs about potential performance, energy levels and opponent's behaviour. While conceptually straightforward, this can quickly lead to intractable problem set-ups; we discuss algorithmic solutions in the next section.

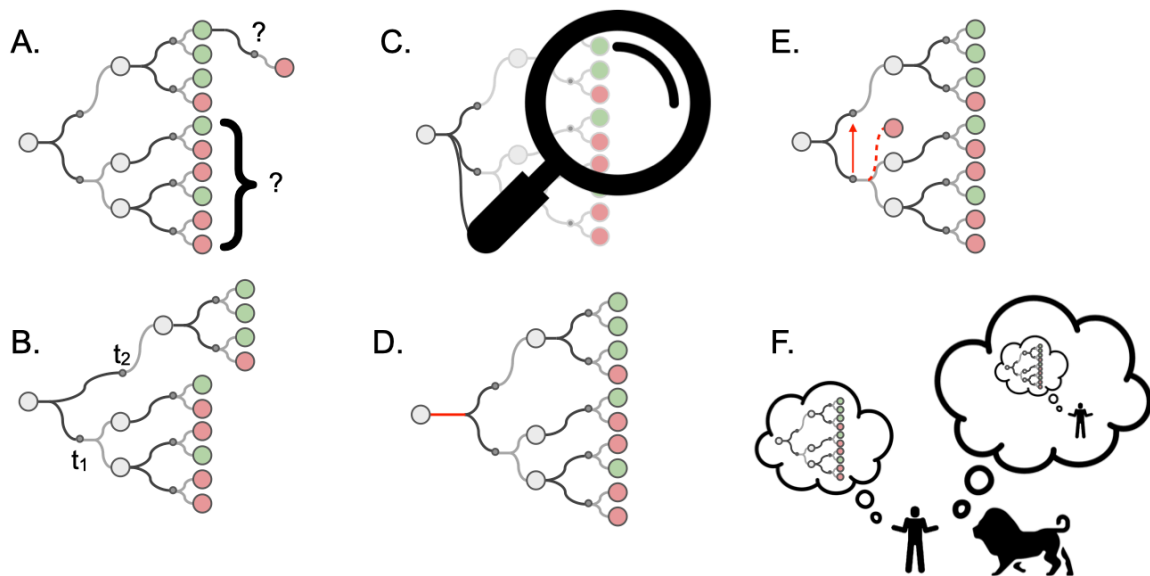


Figure 2: Major computational challenges in finding an optimal defensive action. A. Objective: Balancing short-term lethal outcomes (red circles) and long-term resources (green circles); and how to set the time horizon, for example when short-term favourable state (top green circle) ultimately leads to a negative outcome (e.g. by exhaustion). B. Timing: Putting the decision tree into real time. The two actions in the first state (left circle) need to be taken at different times (for example, escape needs to be started when the opponent is far whereas a fight is only effective when the opponent is near enough). C. Uncertainty in the outcome tree compels information-gathering. Here, this is depicted by a specific action (the magnifying glass) in the first state. D. Uncertainty can also favour robust actions, such as movement trajectories that can be diverted later, illustrated here by the red line segment in the first state. E. Adaptability is required when new information arises during an action. For example, an agent may realise that a specific action will lead to an unforeseen outcome (dashed red line), and revert to a different action (red arrow) that was not part of the initial plan. F. Whilst a prey organism is planning its response, the opponent might form a mental model of the prey and prepare for the prey's most likely action sequence.

Action selection in the face of uncertainty

Next, threat environments are typically imbued with uncertainty on various levels (Bach et al., 2011), which poses several challenges. First, many animals engage in **information-gathering behaviours** (figure 2c) to reduce uncertainty before an initial action (Sih, 1992; Stephens, 2008; Trier et al., 2023). The challenge here is to balance the risks and gains of information gathering (Chittka et al., 2009), and to optimise the type of information that is solicited.

Computationally, dynamical active sampling models (Cassey et al., 2013; Heng et al., 2020; Ozbagci et al., 2021) formalise the inherent trade-off by contrasting the quality of each sample with its internal or external costs, the stakes, and the remaining time.

When uncertainty cannot be resolved or if there is no time to do so, then the next challenge is to select **robust actions** that remain functional in different circumstances (figure 2d). Empirically, this manifests as multi-purpose actions, for example adopting an initial posture that is beneficial for the most dangerous predator location even though somewhat detrimental for other cases (Domenici, 2010; Turesson et al., 2009) or by movement trajectories that allow reaching multiple possible targets (Alhussein and Smith, 2021).

Computationally, proposed strategies are to base behaviour on an explicit distribution of possible states of the environment (Rao, 2010), to plan for the worst possible state of the environment (Weintraub et al., 2020), to consider worse outcomes (Gagne and Dayan, 2022), to avoid actions that entail variability of overall cost (i.e. economic risk) (Nagengast et al., 2010), or to ensure functionality under all possible states of the environment (Weintraub et al., 2020). Some of these strategies were explicitly developed for defensive scenarios, notably in the framework of differential game theory. However, they have been devised for low-dimensional or discrete choice scenarios. In realistic circumstances, the worst state of the environment might not be known even on a phylogenetic level, a fact that can be exploited by "rare enemies", such as tentacled snakes (Catania, 2011). In the face of such ignorance not just about the state of the environment but about its underlying structure, a computational solution is "robust planning", i.e. finding the best policy under the worst possible version of the environment (Buffet and Aberdeen, 2005). Finally, the efficiencies of approximations to an optimal strategy arguably depend on the characteristics of the ecological niche that the agent inhabits (Trimmer et al., 2011).

If new information arrives and resolves uncertainty after an initial action has already been initiated, the agent should adapt its behaviour to take this into account (figure 2e). We discuss the challenges with online adaptation in the next section; here we focus on a priori specification of **adaptable actions**. Empirically, adaptability manifests in many different ways. First, responses can be made adaptable

during the period of decision, preparation, or waiting for the optimal time point, as in ongoing postural preparation and trajectory adaptation before response initiation (Card and Dickinson, 2008; Kimura et al., 2022; Turesson et al., 2009), energisation of locust hindlegs whilst movement trajectory preparation with frontlegs is ongoing (Santer et al., 2005), or adaptation of lizard escape timing to changes in opponent speed (Cooper Jr, 2006). Second, responses can be adaptable whilst already being executed, such as switching from freezing to fleeing or mobbing in fish and birds (Abolins-Abols and Ketterson, 2017; Courter and Ritchison, 2012; Liden et al., 2010), adaptation of hiding movements in locusts (Hassenstein and Hustert, 1999) or escape trajectories in crickets (Sato et al., 2019) and birds (Tätte et al., 2020), and interruption of ongoing escape in humans (Sporrer et al., 2023).

Computationally, adaptability can be achieved by closed-loop control, but also by the specification of (potentially open-loop) multi-phase responses, either with decision points between the phases, as suggested in crabs (Hemmi, 2005; Hemmi and Pfeil, 2010), beetles (Gilbert, 1997), and mice (Shamash et al., 2021), or by an initial open-loop phase followed by adaptable behaviour, as seen in fish (Bullock, 1984). Closed-loop control requires specification of several characteristics. One is the selection of monitorable features of the environment, balanced against the cost of collecting this information. For instance, visual monitoring of a predator during escape might require costly trajectories or head turns; auditory monitoring might be a cheaper but less precise alternative. Another feature is determining adjustable motor plans, such as choosing a speed that provides appropriate manoeuvrability (Howland, 1974; Wynn et al., 2015) or setting fixed safety margins (Hasson et al., 2012). This entails mathematically rich and mechanically challenging trade-offs (Li, 2017). A notion of energy expenditure is also required to avoid exhaustion. Finally, rapid adaptability can be facilitated by parallel planning for, and arbitration between, multiple scenarios, such that cached alternative plans can be rapidly retrieved during action execution (Cisek, 2007; Gallivan et al., 2018, 2015).

Game-theoretic considerations

A final challenge comes from **game-theoretic considerations** in agent-opponent interaction (figure 2f). Empirical examples are deceptive signalling (Caro, 2014; Cooper, 1998) and protean, i.e. unpredictable, behaviour (Evans et al., 2019; Humphries and Driver, 1970). The latter is particularly relevant if a predator moves faster than the prey, and can save the latter from otherwise certain death even though further reducing speed. Empirically, unpredictable tactics are used in the selection (Arnott et al., 1999; Domenici et al., 2008) and execution (Domenici and Hale, 2019; Jornod and Roche, 2015) of behaviour in many species and can depend on threat characteristics (Herbert-Read et al., 2017; Storms et al., 2019). Opponents, of course, can use the same strategies, which engenders a game theoretic arms race between prey and predator, each attempting to outwit and deceive the

other (Boesch, 2002; Catania, 2011; Hein et al., 2020). The problem space is even larger in gregarious species in which groups of individuals jointly decide on defensive behaviour (Roberts, 1996).

Computationally, unpredictability can arise from stochastic actions (Peterson et al., 2021) or deliberate deception. Stochastic action-selection has strong theoretical foundations (Hausken and Levitin, 2009; Humphries and Driver, 1970) and practical advantages, including low computational burden and broad applicability. More sophisticated strategies can be found using differential game theory tools under incomplete information (Weintraub et al., 2020), or by adapting POMDP models to multi-agent settings (Gmytrasiewicz and Doshi, 2005), although practical solutions might only ever be approximately correct.

Overall, the computational challenges outlined can be grouped into two categories. One encompasses strategic choices that partly depend on the biophysics of the body and the capacities of the neural system. These can be solved prior to a threat encounter, and might well be hard-wired from birth or during development: determining action objectives and time horizons, robustness and adaptability, and a broad attitude towards opponents. Within these strategic settings sit other challenges that comprise tactical decisions, many of which must be flexible from situation to situation: computing optimal response timing and actual robust plans, adapting behaviour online, and implementing appropriate game-theoretic responding. How this could be achieved under time pressure is the topic of the next section.

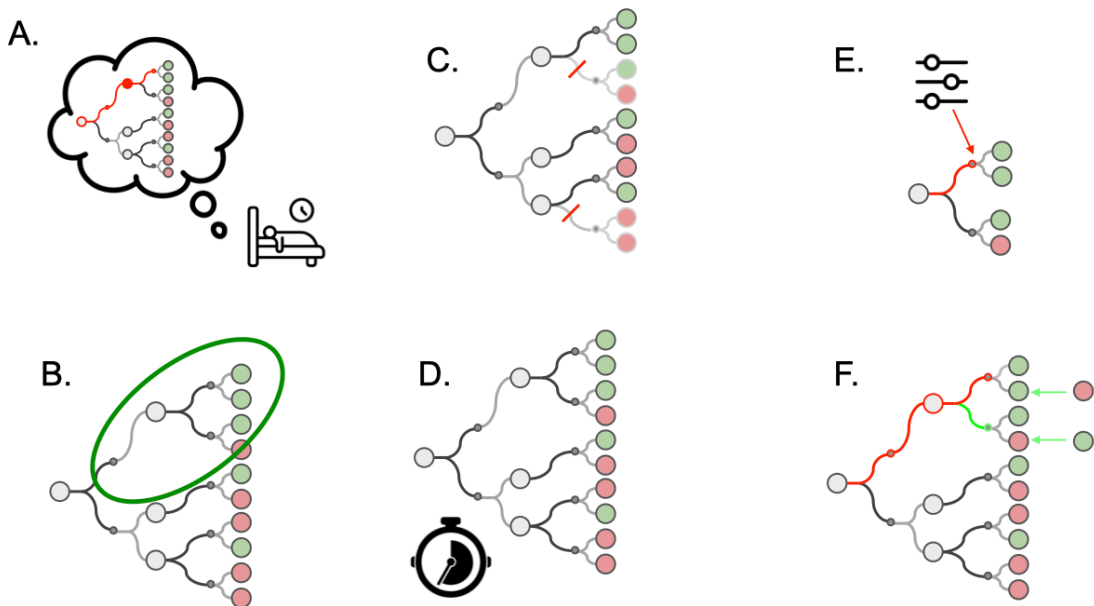
The algorithmic level: Selecting the defensive response

Even if it is possible to specify an optimal strategy, substantial further demands are imposed by the need to determine the actual responses and adapt them online. Of course, an agent can opt to respond with reflexive input-output relationships to a specific set of sensory signals without analysing the remaining environment, such as reflexes (Yeomans et al., 2002; Bach, 2015a; Domenici and Hale, 2019; LeDoux and Daw, 2018), and other relatively stereotypical behaviours (Bach and Dayan, 2017; Marras and Domenici, 2013). An agent might also prioritise certain sensory patterns for detailed perceptual and mnemonic processing (Bach et al., 2014; LoBue et al., 2014; Schmidt-Daffy, 2011; You and Li, 2016). Yet another reflexive strategy is to seek out or avoid attractive and repulsive parts of the environment respectively (Shi et al., 2017; Woodbury, 1986). Extending such ideas, Hein et al. (2020) provide a revealing analysis of behavioural algorithms: a small set of reflexive input-output relationships that provide adaptability at run time, and are relatively straightforward to select between.

403 However, there is by now much evidence that even seemingly simplistic escape-to-shelter behaviour
404 in mammals depend on a detailed assessment of the environment, including the trajectory and
405 identity of the opponent (Evans et al., 2019; Sporrer et al., 2023). This results in a tree of possible
406 actions and responses that expands into the future, which necessitates some forms of planning
407 (Mattar and Lengyel, 2022). Reflective algorithms require resources and therefore inflict their own
408 challenges. Indeed, compromises between optimal and satisficing (Simon, 1955) will typically be
409 necessary – the confounding of Marr & Poggio (1977)’s levels mentioned above.

410 Reducing complexity

411 A primary challenge of action planning in many situations involves **rapidity**. Regardless of the agent's
412 current activity or level of preparation, it must often react promptly to ensure survival (Walker et al.,
413 2005). Thus, ongoing behaviour must be interruptible to enable the selection and execution of an
414 appropriate escape response, and these new actions must be selected quickly. Moreover, windows of
415 opportunity for some – but not others – of the available actions might already close during the decision
416 process (Reynaud et al., 2020). While the neural system has ways to accelerate decision processes,
417 such as invigoration (Manohar et al., 2015) and specialised low-latency circuits (Bullock, 1984), rapidity
418 will often require **appropriate reduction of the complexity** (but therefore also, quality) of online
419 planning (figure 3). More generally, this is a necessity even without time pressure: agents in real-world
420 scenarios operate in open environments with near-infinite possibilities that all but preclude the
421 application of classical near-optimal algorithms such as exhaustive tree search (Gordon et al., 2021)
422 even given a focus on the worst cases (e.g., Lecarpentier & Rachelson, 2019).



425

426 **Figure 3: Major algorithmic challenges and their proposed solutions.** Rapidity and the
 427 multidimensionality of naturalistic behaviour require complexity reductions. Five groups of strategies
 428 are proposed: A. Pre-planning during restful states. B. Hierarchical planning such as chunking
 429 (representing an entire part of the decision tree as one option). C. Sparsifying the decision tree, e.g.
 430 by pruning (shown here) or sparse sampling. D. Anytime algorithms ensure a good-enough decision
 431 can be read out at any time during the decision process, for example when time is running out. E.
 432 Factorisation. An agent might first decide on an action, and separately, on its implementation
 433 parameters such as optimal timing. F. Continuous updating. As new information arrives (green
 434 arrows), an initial plan (red trace) needs to be updated (green trace), preferably without having to re-
 435 compute the entire action sequence.

436 Many algorithmic methods have been suggested to this end (albeit mostly in appetitive rather than
 437 aversive domains), and likely are collectively employed and combined given the difficulty and
 438 importance of the problem (Collins & Cockburn, 2020; Feher da Silva & Hare, 2020). Here, we group
 439 them into five categories: offline (pre-)planning, hierarchical planning, sparse planning, anytime
 440 planning, and factorisation.

441 First, agents may take advantage of epochs before a threat appears (Mobbs et al., 2020) or has yet to
 442 attack (Roelofs & Dayan, 2023) to **pre-plan** or adjust counterfactual strategies (figure 3a). Periods of
 443 quiet wakefulness (e.g., during consumption of food in a currently safe environment or embedded in

a larger group (Roberts, 1996)) or sleep (e.g., after spending time in an environment that is likely to be revisited) have been suggested as the ideal time during which to leverage a world model in order to build an effective policy that can be executed reflexively when the need arises (Sutton, 1990; Momennejad et al., 2018; Mattar & Daw, 2018). Mice appear to memorize spatial subgoals during exploration, and escape towards them even when the environment has changed, indicating pre-planning (Shamash et al., 2021). Similarly, complex escape routes in mice improve with prolonged exposure to the spatial context (Claudi et al., 2022). There are various versions of offline-trainable reflexive policies, including behavioural habits (Dickinson & Balleine, 1991; Gershman et al., 2016; Gershman, 2020) and model-free methods from reinforcement learning (Daw et al., 2005). To address capacity limits in the representation of ensuing policies, agent might use policy compression (Lai and Gershman, 2021).

Next, **hierarchical** approaches to simplify planning (Botvinick, 2012; Janssen et al., 2022) have been proposed on a cognitive, neural and motor control levels. On a cognitive level, chunking (figure 3b) could help roll out entire action sequences in predictable environments (Wu et al., 2023), which may manifest as inflexible habits (Balleine and Dezfouli, 2019) to achieve subgoals (Q. J. Huys et al., 2015). Subgoals are particularly useful if the environment contains a hierarchical structure, in which action planning could progress from coarse to fine levels of detail over time (Tomov et al., 2020). Neural and neurally-inspired theories have posited that that planning not only takes place in hierarchical temporal order, but simultaneously in a hierarchically structured (neural) planning system. Here, higher levels of a neural processing hierarchy set subgoals for lower levels, which are then achieved rather autonomously (Friston, 2010; Pezzulo and Cisek, 2016). Subgoals that involve reaching desired future states can usefully be formalised as behavioural affordances (Pezzulo and Cisek, 2016). The motor system faces the related problem that the abundance of degrees of freedom complicates forward planning. Similar hierarchical principles offer a solution, with the selection of motor goals being partially independent of the planning for their implementation (Merel et al., 2019) and with the latter again proceeding somewhat autonomously, for instance by changing the reference points of reflex arcs (Latash, 2021). This would then involve the periphery outside the central nervous system in the planning.

Next, given the impossibility of optimal planning, some of the many forms of satisficing (Simon, 1955) and resource- (Lieder & Griffiths, 2020) and bounded- (Devaine et al., 2014; Prystawski et al., 2022) rationality will apply to **sparsify** the decision tree (figure 3c) and reduce its complexity (Huys et al., 2015). A full tree could be pruned to remove costly and risky spatiotemporal areas (Guo et al., 2022); the pruning could be adaptive (Snider et al., 2015), and the parts removed could be replaced by approximations produced by simpler, e.g., model-free methods (Huys et al., 2015; Ott et al., 2020)

according to the available time (Keramati et al., 2016). Alternatively, only parts of the tree could be built in the first place, as in stochastic scenario-exploration—a form of sparse sampling—(Kearns et al., 2002, 1999) focusing on the more probable (Ye et al., 2017) scenarios, the ones with higher expected returns (Mern et al., 2021) or higher danger (Lieder et al., 2014; Luo et al., 2019). Computationally simplified sampling explorations such as Monte Carlo tree search (Coulom, 2007) unlock the use of POMDP for real-world planning tasks (Silver & Veness, 2010; Bai et al., 2015), bridging the gap between this principled solution to planning under uncertainty with realistic computational time (Hay et al., 2012; Lieder and Griffiths, 2020). These methods are more effective if their operations can occur in parallel (Cai et al., 2021). However, such a solution critically depends on the agent's computational capacities, with shared representations acting as a bottleneck (Petri et al., 2021), and it has been suggested that Monte Carlo tree search would be too slow for common escape scenarios in rodents (Espinosa et al., 2022).

An additional advantage of sample-based methods such as Monte Carlo tree search is that they are **anytime algorithms** (Tonola et al., 2023, 2021; Zilberstein, 1996), which provide (progressively more useful) solutions at any time during the decision process, rather than having to run to completion (figure 3d). Biological examples of anytime algorithms include diffusion decision models (Ratcliff et al., 2016), in which decision bounds can be collapsed at any time (Gold and Stocker, 2017; Shadlen and Kiani, 2013; Shea-Brown et al., 2008). Agents could also use leaky evidence integration in conjunction with contracting bounds as in the urgency-gating model to account for continuously changing environmental states (Thura et al., 2012; Thura and Cisek, 2017; Puelma Touzel et al., 2022), or otherwise time-varying bounds to account for dynamic opportunities (Barendregt et al., 2022). In addition, active sampling can be integrated into these models (Ozbagci et al., 2021). There is some evidence for the existence of drift-diffusion processes in biological agents exposed to threat (Evans et al., 2018; Lee et al., 2013), and evidence from simple motor tasks that these processes feed into movement preparation already before a decision threshold is reached (Selen et al., 2012). Alternatively, decision algorithms that compute action outcomes sequentially can return solutions at any time, while satisficing thresholds can ensure some minimal quality, albeit at the expense of under- or over-patience. There is also a possibility of the adaptive deployment of cognitive control to combine information from multiple algorithms with different completion times (Yeung and Summerfield, 2012; Shenhav et al., 2013; Lieder et al., 2018; Frömer et al., 2021).

Finally, although there are various challenges (Cisek, 2007; Gallivan et al., 2018) to a classical serial view in which action selection always precedes the specification of the execution parameters, there are circumstances under which parameters of the action can be at least partially **factorised** (figure 3e), and optimised independently. This simplifies the decision problem (Gallivan et al., 2018). An

example is response timing, which is of particular importance due to its strong biological (neural and biomechanical) constraints. Response timing can be thought of as a parameter of the optimal response, but could also be optimised in separate planning schemes after a response has been selected (Dayan, 2012; Bach, 2015b, 2017).

Continuous adjustment

Once an initial plan is made, this plan should adapt to the **continuous delivery of new information while the action is carried out** (figure 3f). This is a classic challenge in embodied decision-making (Cisek and Pastor-Bernier, 2014; Gordon et al., 2021) and non-trivial, as many decision algorithms need time to complete and cannot update continuously.

Computationally, continuous control can be implemented by anytime algorithms including biologically motivated drift-diffusion models which formalise evidence accumulation beyond an initial choice (Stone et al., 2022) and can implement action changes after action initiation (Resulaj et al., 2009). From an engineering perspective, there are several computational strategies of interest: continual or online planning (Brenner and Nebel, 2006; Majumdar and Tedrake, 2013; Sinha et al., 2020), replanning under limited knowledge (Brafman and Shani, 2012), anytime replanning (Tonola et al., 2021), and re-use/repair of previous plans (Fox et al., 2006; Guzman et al., 2014; Scala et al., 2015; Tonola et al., 2023), but few of these have been assessed in terms of biological plausibility (although see Piray and Daw, 2021). In environments with discontinuous states, change point detection algorithms can help determining the time point at which current plans are inappropriate (Wilson et al., 2010). Also, multiple actions can be planned in parallel, reducing the complexity of the online decisions problem to switching between (Cisek, 2007; Gallivan et al., 2018, 2015). For example, the affordance competition hypothesis posits that (at least mammal) agents maintain a continuous representation of currently available motor plans (Cisek, 2007). However, such strategies are limited by physical (Marois and Ivanoff, 2005; Tombu et al., 2011) and informational bottlenecks (Koechlin and Hyafil, 2007; Musslick and Cohen, 2021) that would likely differ across species and environments.

To summarize, the primary algorithmic challenges of critical intelligence involve an effective reduction of complexity and continuous adaptation in the face of uncertainty. The reduction of complexity, in particular, poses a danger as it goes against the demands of optimality. Instead of aiming for the best response, the agent will seek to find a satisfactory response based on certain criteria. It seems plausible to assume that the choice of algorithm(s) is hard-wired at birth or during development; but given the evident sophistication even of the simplest defensive behaviours, many of the actual computations will have to be performed online. Thus, our analysis shifts the focus from delineating a hard-wired set of defensive behaviours to highlighting a hard-wired set of algorithms to compute

behaviour. Neuroscience is only now equipped with the technology to investigate naturalistic behaviour in such circumstances.

A final aspect of adaptability is offline adaptation – after a threat encounter has been survived. In the next section we discuss what the potential strategies are to this end.

Learning without direct experiences

To improve survival odds, defensive responses should be adapted to those environments and predators that an animal might encounter, and such adaptation occurs on multiple time scales (Lima and Dill, 1990; Sporrer et al., 2023). Here, we focus on how an agent can harness threat encounters in order to learn better survival strategies prior to the next encounter. Critical to this process, agents must adapt their response without directly experiencing complete failure.

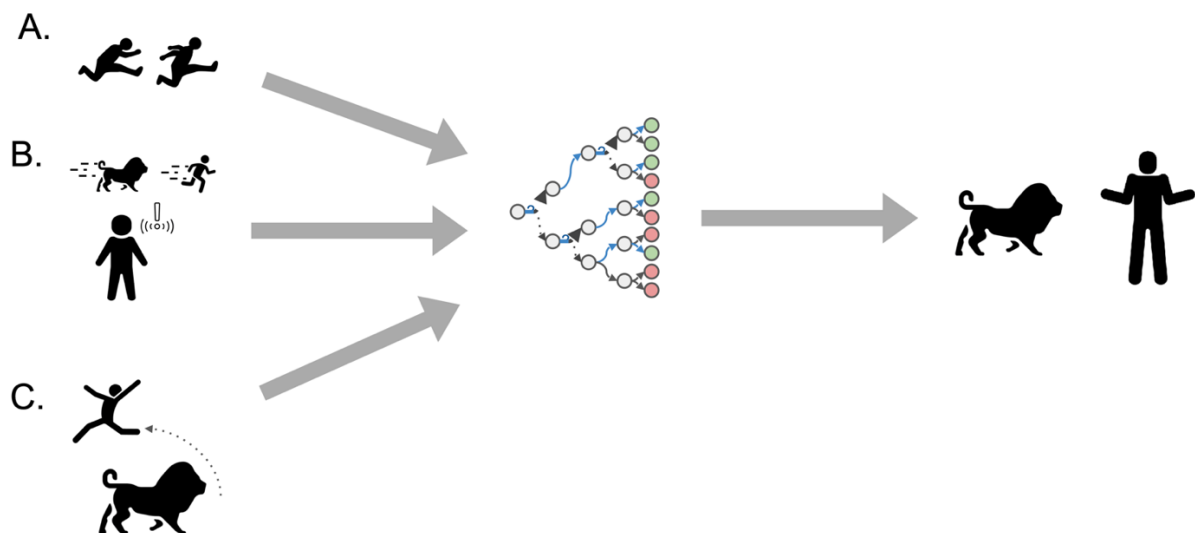


Figure 4: Learning without direct experience. Three ways to learn without direct experience: A. Rough-and-tumble play is observed in many mammal species especially during adolescence. B. Observing conspecifics in the same situations. C. Learning from near-catch experiences. In each case, a mental model of the environment is constructed from the experience, and then adapted and generalised such as to be applicable in an actual threat encounter.

The first method of learning without direct experience is **generalising from safe experiences** (figure 4a). In the absence of actual deadly experiences, an agent can emulate them. For instance, by engaging in play fights with its peers, such as rough-and-tumble play in many mammals (Fagen, 1976; Spinka et

al., 2001; Weller et al., 2020), agents may practice defense against conspecifics. Similarly, mice appear to learn escape paths from spontaneous escape attempts in the absence of threat (Shamash et al., 2021).

Computationally, this can be achieved through two complementary strategies, transferring knowledge from one situation to another and broadening the range of experiences to draw from. In the former, internal models are fine-tuned to a specific situation, like wrestling with peers, and adjusted to account for another, different but comparable situation, like fighting an opponent. Cognitive maps and computational models of the hippocampal and entorhinal systems are interesting candidates for such a process (Tolman, 1948; O'Keefe et al., 1978; Eichenbaum, 2015; Whittington et al., 2020), alongside less biological approaches for artificial systems (Weiss et al., 2016; Niu et al., 2020). The latter strategy consists of seeking a variety of training scenarios, encompassing various opponents, tactics, and environmental conditions (Bell et al., 2009; Salvanes et al., 2013; Tetzlaff et al., 2019). Curiosity-driven algorithms can achieve such behaviour through intrinsic goals and self-evaluation processes (Gottlieb et al., 2013; Ten et al., 2021), yet they are typically not evaluated in threat scenarios. Finally, agents can also generalise from one threat to another (Anson and Dickman, 2013; Griffin et al., 2001).

The second method of learning without direct experiences is **learning from observation** (figure 4b). Observing an interaction between a conspecific and its opponent, without any direct interaction, provides critical information about the opponent's approach strategy and any potential deceptive tactics. In turn, this information can fine-tune one's own defensive strategies, for instance, by paying more attention to opponents' potential hideouts. Behaviourally, it is well-known that various mammals use this type of information, evidenced for example by conditioned responses after observing Pavlovian conditioning in a conspecific, something that requires anterior cingulate to amygdala projections in rodents and probably also in humans (Olsson et al., 2020; Smith et al., 2021). However, the challenge here is assessing the other agent's experiences indirectly. During observation, the observer might have information that it does not when threatened itself (e.g. due to a different view angle), and also the threatened agent might have information that the observer does not when threatened itself (e.g. due to different initial threat responses). Thus, the challenge is to adjust defensive strategies based on indirect scenarios analogous to actual experiences of the threat. Interestingly, besides observation, parental influences might also be transmitted more directly, i.e. via hormonal or epigenetic means (Atherton and McCormick, 2020).

Computational approaches to this problem have been proposed in robotics and AI research (Lu et al., 2018). Through imitation of their parents and peers, for instance, an agent could emulate and train defensive responses, while minimising risks. Complementarily, an agent can simulate what could have

alternatively happened during an encounter (Mattar and Daw, 2018; Sutton, 1990), potentially using replay to learn from memorised observations (Panoz-Brown et al., 2018), and adjust its future defensive response accordingly, for instance by strengthening a reflexive response to stimuli.

The third method of learning without direct experiences is **learning from near-catch experiences** (figure 4c). Almost being caught is the closest to a negative example an agent can get during its lifetime. More generally an agent can extend its notion of loss to adverse situations that could have led to death, such as engaging in a fight, being hurt or simply being too close to a predator. However, unlike classical reward-maximisation scenarios, adjustment steps must be chosen carefully to avoid any worse outcomes.

Computationally, an agent can rely on constrained optimisation algorithms' mechanisms. Variations of a strategy can then be tested within the realm of safe and known responses. After repeated approaches, different species of grasshoppers for instance, gradually adjust their escapes based on their species' characteristics, either strengthening their jump or hastening their flight (Bateman and Fleming, 2014). The difficulty is then to adjust the intensity of the constraint with the risk of a future encounter. A dramatic near-catch would call for a strong response change. Again, offline simulation processes, like the DYNA algorithm (Mattar and Daw, 2018; Sutton, 1990), could provide such estimates by gradually widening the range of simulated scenarios.

To summarise, agents can employ several strategies to learn from indirect experience. Given ubiquitous observation in mammals of rough-and-tumble play, vicarious learning, and adaptation immediately after threat encounters, it is likely that all of them are used to some extent. However, it is less clear how they might be efficiently and effectively implemented on an algorithmic or neural level. This resonates with our general theme, which we expand on in the next section.

Towards understanding neural implementation

The neuroscience of defensive behaviour has traditionally focused on finding neural circuits that implement protective and defensive reflexes (Bullock, 1984; Davis, 2006; Walters et al., 1981; Yeomans et al., 2002), or stereotypical defensive behaviours such as freezing, undirected escape, passive avoidance and the choice between them, recently using sophisticated circuit-mapping techniques in mice (e.g., Herry et al., 2008; Fadok et al., 2017; Kennedy et al., 2020) and flies (Duistermars et al., 2018; Gibson et al., 2015; Mohammad et al., 2016). Because these types of behaviours take a relatively small part of the sensory environment as an input, they can be elicited in well-controlled laboratory circumstances, sometimes even with head-fixed animals. While this has generated impressive insights, these fall short of explaining the sophistication of defensive behaviour

in mammals in more naturalistic settings (Domenici and Hale, 2019; Evans et al., 2019; Sporrer et al., 2023), including the many continuous parameters of behaviours. We argue that understanding this large action space requires a novel approach; both in terms of the research paradigms, and in the level of theory. Our analysis of the environmental demands and the potential computational and algorithmic solutions highlights what we should be looking for in the neural system: not circuits that implement particular behaviours, but rather circuits that implement particular algorithms, which flexibly decide between various behaviours and continuously compute and update their parameters. Because this research programme is only just beginning, we can but give a brief overview of emerging ideas in this field.

Some literature has addressed relatively threat-specific circuits in mice that implement partly pre-planned escape to shelter (Shamash et al., 2021) from a rapidly expanding overhead disc (Evans et al., 2018). This research has identified circuits for maintaining important decision variables. First, the spatial target, i.e. shelter orientation, is tonically represented in specific neurons in superior colliculus and retrosplenial cortex (Campagner et al., 2023). Second, a race-to-threshold mechanism in superior colliculus has been described for eliciting escape, regardless whether elicited by visual or auditory input (Evans et al., 2018), with the threshold for this mechanism implemented in dorsal periaqueductal gray (Evans et al., 2018). This threshold is under modulatory control by GABAergic interneurons (Stempel et al., 2024), while the accumulation process in superior colliculus is modulated by input from ventral lateral geniculate nucleus, which signals contextual variables (Fratzl et al., 2021). Taken together, this suggests a dedicated circuit for flexible escape responses, which appears to integrate pre-planning of details, an evidence accumulation process for elicitation, and lack of a detailed environment model at run-time in what appears as open-loop phases.

Regarding the neural implementation of goal-directed behaviours, there is a scarcity of research in defensive situations. More generally, cortical and hippocampal replay (Olafsdottir et al., 2018; Panoz-Brown et al., 2018; Pfeiffer and Foster, 2013), is a suitable candidate for offline preparation and learning (Mattar and Daw, 2018; Gagne and Dayan, 2022). Hippocampal replay might also be a mechanism of forward simulation of a decision tree. For example, replay of spatial paths has been observed in relation to the future locations that an animal will visit (Ferbinteanu and Shapiro, 2003). Potentially indicating the same mechanism, in the vicinity of previously fear-conditioned locations, replay of paths towards these locations (which are behaviourally avoided) has been shown (Wu et al., 2017).

However, these are suggestions based on scarce evidence or through analogies. To make headway, it will be important to collect neural observations in freely moving animals and humans; technology that has recently become available. Given that critical neural computations are performed in distributed,

partially-modular, populations (Basu et al., 2021; Churchland and Shenoy, 2024), a particular focus on observing large numbers of neurons appears warranted.

Conclusion

Investigating the sophistication of defensive behaviour entails experiments that actually afford such sophistication – requiring that we provide agents with complex, naturalistic environments, in which they have many and continuous action choices. The technology for such experiments has only recently become available, marking the beginning of a long and exciting journey into the realm of critical intelligence in biological agents.

As a backdrop, our review focused on abstract, computational and algorithmic, challenges of defensive behaviour. Our analysis highlights two main points. First, as generally in naturalistic environments, it is practically impossible to compute optimal behaviour, due to the large state and action spaces. However, second, unlike more general instantiations of embodied decision-making, the problem setting also constrains the algorithms that can be used to compute behaviour. For just one instance, since exploration is not an option in the face of lethal threat, the onus is on robust actions and adaptability, both in real time and after the encounter. A central theme emerges: approximating ideal responses through heuristics that balance safety and practicality. With this review, we intend to provide a roadmap for the empirical programme of characterising these heuristics: first, by narrowing down the focus of neural experiments; secondly to provide benchmark solutions from the wider machine learning and computational neuroscience literature against which neural data and simulations can be evaluated.

Importantly, distinguishing between different heuristics from observations alone can be difficult (Mattar and Lengyel, 2022). For example, different reflective, tree-search, algorithms converge to similar solutions, given sufficient time or adequate state spaces (a many-to-one mapping). At the same time, the interplay between a specific algorithm and different reflexive strategies could obscure some of its specific characteristics (a one-to-many mapping). In general, there are two ways of solving this conundrum. The first is to leverage neural observations, in order to find those algorithms that are actually implemented in the neural system (Mattar and Lengyel, 2022). The caveat here is that we lack ground truth as to how the brain computes behaviourally relevant outputs and therefore, which neural features to record and analyse. Important discoveries have often been made in rather constrained experiments, where the instantaneous state of the system is more easily defined externally. In naturalistic environments, which are required for this research programme, it will often

be unclear a priori what the neural system's goals are at a specific time. Knowing (or hypothesising) a specific algorithm that it might use can then guide scientific discovery.

The second possibility for distinguishing candidate algorithms is to use behavioural perturbation experiments: systematically creating situations in which only specific algorithms fail, while others succeed. This can provide a broad overview as to which features an algorithm uses. Historical examples are detour tasks (Tolman, 1948) or devaluation experiments (Dickinson and Balleine, 1994). Similar ideas have been used to investigate the structure of algorithms for defensive behaviour (Bach, 2017; Sporrer et al., 2023). This has revealed, for example, that while reflexes and reflexive behaviour do exist across the animal kingdom including humans, they do not account for the majority of observed defensive behaviour in many mammalian species (Evans et al., 2019; Sporrer et al., 2023). In textbox 1, we summarise conceptual suggestions as to how to investigate some of the mechanisms proposed in this article. Importantly, for simple defensive responses it is well-known that different species use different algorithms, and there is no reason why this should not be the case for more complex behaviour. Thus, knowledge of the ecological niche in which a species is operating might help elucidate suitable candidate algorithms (Kavaliers and Choleris, 2001).

It is important not to forget additional constraints on the planning and execution of behaviour. For instance, the reason why certain defensive responses are non-interruptible (Kimura et al., 2022) could be biomechanical (ballistic movements) and/or neural (open-loop control). Equally, biological agents have to navigate continuous state and action spaces – an area which is poorly explored over and above rather simple motor actions. There is a suggestion that behaviour is altogether composed of discrete elements, such as motor primitives on a fine-grained level (Chiovetto et al., 2022), or behaviour syllables on a coarser level (Wiltschko et al., 2015). Such a hierarchical scheme might enable largely discrete planning, but the evidence that such discrete elements are used for action planning is limited.

One potential source of computational and algorithmic ideas for critical intelligence comes from the broad fields of real-time systems engineering (Kopetz and Steiner, 2022) and what are known as cyber-physical systems (Liu et al., 2017). These directly address the engineering challenges associated with operating effectively and safely within the constraints of realized physical systems, and offer an extensive body of theory and practice that could directly inform our understanding of biological systems.

In conclusion, this article provides a roadmap for advancing critical intelligence research in naturalistic environments, by identifying the most challenging computational and algorithmic issues and reviewing potential statistical solutions. To bridge the gaps in our understanding of their neural implementation, it is essential to conduct naturalistic neuroscience experiments. Although these

experiments are more complex to control than traditional methods, they provide invaluable insights into crucial natural processes that extend beyond current computational theories that were based on constrained settings.

Glossary

State: Current situation of the agent, in terms of the external environment, internal environment, and its past history of interaction with these.

Policy: a mapping from states to actions, can be deterministic or probabilistic.

Reflexive algorithm: a direct mapping from external states (or continuous control variables) to actions

Behavioural algorithms: a set of reflexive algorithms, equipped with a reflective meta-algorithm that selects between them.

Reflective algorithm: a decision-making strategy based on simulations of a model of the environment, also termed "planning".

Planning: the use of reflective algorithms to specify one's actions.

Model-free reinforcement learning: learning based on increasing the probability of rewarded actions.

Model-based reinforcement learning: learning a model of the environment.

Partially observable Markov decision processes (POMDP): a stochastic process in which the transition of each state to another state objectively only depends on the current states and the actions of the agent, but the current state is not fully known to the agent and needs to be inferred.

Utility: a real-valued "common currency" to value actions, and thus to balance an agent's different objectives used on economic decision-making theories.

Loss: negative utility, used in many statistical decision theories.

Closed-loop control: controlling a dynamical system based on feedback from the system's own output.

Open-loop control: controlling a dynamical system without feedback from the system's output.

Satisficing: This portmanteau of "satisfy" and "suffice" describes searching the decision space until a "good enough" outcome is identified. The term was coined for business situations to which the management can respond by "finding optimum solutions for a simplified world, or by finding satisfactory solutions for a more realistic world" (Simon, 1979).

References

- Abolins-Abols, M., Ketterson, E.D., 2017. Condition explains individual variation in mobbing behavior. *Ethology* 123, 495–502. <https://doi.org/10.1111/eth.12625>
- Alhussein, L., Smith, M.A., 2021. Motor planning under uncertainty. *eLife* 10, e67019. <https://doi.org/10.7554/eLife.67019>
- Anson, J.R., Dickman, C.R., 2013. Behavioral responses of native prey to disparate predators: naiveté and predator recognition. *Oecologia* 171, 367–377. <https://doi.org/10.1007/s00442-012-2424-7>
- Arnott, S.A., Neil, D.M., Ansell, A.D., 1999. Escape trajectories of the brown shrimp *Crangon crangon*, and a theoretical consideration of initial escape angles from predators. *J. Exp. Biol.* 202, 193–209. <https://doi.org/10.1242/jeb.202.2.193>
- Atherton, J.A., McCormick, M.I., 2020. Parents know best: transgenerational predator recognition through parental effects. *PeerJ* 8, e9340. <https://doi.org/10.7717/peerj.9340>
- Bach, D.R., 2017. The cognitive architecture of anxiety-like behavioral inhibition. *J Exp Psychol Hum Percept Perform* 43, 18–29. <https://doi.org/10.1037/xhp0000282>
- Bach, D.R., 2015a. A cost minimisation and Bayesian inference model predicts startle reflex modulation across species. *J Theor Biol* 370, 53–60. <https://doi.org/10.1016/j.jtbi.2015.01.031>
- Bach, D.R., 2015b. Anxiety-Like Behavioural Inhibition Is Normative under Environmental Threat-Reward Correlations. *PLoS Comput Biol* 11, e1004646. <https://doi.org/10.1371/journal.pcbi.1004646>
- Bach, D.R., Dayan, P., 2017. Algorithms for survival: a comparative perspective on emotions. *Nat. Rev. Neurosci.* 18, 311–319. <https://doi.org/10.1038/nrn.2017.35>
- Bach, D.R., Hulme, O., Penny, W.D., Dolan, R.J., 2011. The known unknowns: neural representation of second-order uncertainty, and ambiguity. *J. Neurosci.*
- Bach, D.R., Schmidt-Daffy, M., Dolan, R.J., 2014. Facial expression influences face identity recognition during the attentional blink. *Emot. Wash. DC* 14, 1007–1013. <https://doi.org/10.1037/a0037945>
- Bai, H., Cai, S., Ye, N., Hsu, D., Lee, W.S., 2015. Intention-aware online POMDP planning for autonomous driving in a crowd, in: 2015 IEEE International Conference on Robotics and Automation (Icra). IEEE, pp. 454–460.
- Balleine, B.W., Dezfouli, A., 2019. Hierarchical Action Control: Adaptive Collaboration Between Actions and Habits. *Front. Psychol.* 10.

786 Barendregt, N.W., Gold, J.I., Josić, K., Kilpatrick, Z.P., 2022. Normative decision rules in changing
787 environments. *eLife* 11, e79824. <https://doi.org/10.7554/eLife.79824>

788 Basu, R., Gebauer, R., Herfurth, T., Kolb, S., Golipour, Z., Tchumatchenko, T., Ito, H.T., 2021. The
789 orbitofrontal cortex maps future navigational goals. *Nature* 599, 449–452.
790 <https://doi.org/10.1038/s41586-021-04042-9>

791 Bateman, P.W., Fleming, P.A., 2014. Switching to Plan B: changes in the escape tactics of two
792 grasshopper species (Acrididae: Orthoptera) in response to repeated predatory approaches. *Behav.*
793 *Ecol. Sociobiol.* 68, 457–465. <https://doi.org/10.1007/s00265-013-1660-0>

794 Bell, J. a., Livesey, P. j., Meyer, J. f., 2009. Environmental enrichment influences survival rate and
795 enhances exploration and learning but produces variable responses to the radial maze in old rats. *Dev.*
796 *Psychobiol.* 51, 564–578. <https://doi.org/10.1002/dev.20394>

797 Berger, J.O., 2013. Statistical decision theory and Bayesian analysis.

798 Bernoulli, D., 1738. Specimen theoriae novae de mensura sortis. *Comment. Acad. Sci. Imp.*
799 *Petropolitanae*.

800 Bialek, W., 2022. On the dimensionality of behavior. *Proc. Natl. Acad. Sci.* 119, e2021860119.
801 <https://doi.org/10.1073/pnas.2021860119>

802 Boesch, C., 2002. Cooperative hunting roles among taï chimpanzees. *Hum. Nat.* 13, 27–46.
803 <https://doi.org/10.1007/s12110-002-1013-6>

804 Botvinick, M.M., 2012. Hierarchical reinforcement learning and decision making. *Curr. Opin.*
805 *Neurobiol.* 22, 956–962. <https://doi.org/10.1016/j.conb.2012.05.008>

806 Brafman, R.I., Shani, G., 2012. Replanning in Domains with Partial Information and Sensing Actions. *J.*
807 *Artif. Intell. Res.* 45, 565–600. <https://doi.org/10.1613/jair.3711>

808 Brenner, M., Nebel, B., 2006. Continual planning and acting in dynamic multiagent environments, in:
809 *Proceedings of the 2006 International Symposium on Practical Cognitive Agents and Robots, PCAR '06.*
810 *Association for Computing Machinery, New York, NY, USA, pp. 15–26.*
811 <https://doi.org/10.1145/1232425.1232431>

812 Brette, R., 2019. Is coding a relevant metaphor for the brain? *Behav. Brain Sci.* 42, e215.
813 <https://doi.org/10.1017/S0140525X19000049>

814 Brookes, J., Hall, S., Fröhholz, S., Bach, D.R., 2023. Immersive VR for investigating threat avoidance:
815 The VRthreat toolkit for Unity. *Behav. Res. Methods.* <https://doi.org/10.3758/s13428-023-02241-y>

816 Buffet, O., Aberdeen, D., 2005. Robust planning with (L)RTDP, in: Proceedings of the 19th International
817 Joint Conference on Artificial Intelligence, IJCAI'05. Morgan Kaufmann Publishers Inc., San Francisco,
818 CA, USA, pp. 1214–1219.

819 Bullock, T.H., 1984. Comparative neuroethology of startle, rapid escape, and giant fiber-mediated
820 responses, in: *Neural Mechanisms of Startle Behavior*. Springer, pp. 1–13.

821 Cai, P., Luo, Y., Hsu, D., Lee, W.S., 2021. HyP-DESPOT: A hybrid parallel algorithm for online planning
822 under uncertainty. *Int. J. Robot. Res.* 40, 558–573. <https://doi.org/10.1177/0278364920937074>

823 Campagner, D., Vale, R., Tan, Y.L., Iordanidou, P., Pavón Arocas, O., Claudi, F., Stempel, A.V.,
824 Keshavarzi, S., Petersen, R.S., Margrie, T.W., Branco, T., 2023. A cortico-collicular circuit for orienting
825 to shelter during escape. *Nature* 613, 111–119. <https://doi.org/10.1038/s41586-022-05553-9>

826 Card, G., Dickinson, M.H., 2008. Visually Mediated Motor Planning in the Escape Response of
827 *Drosophila*. *Curr. Biol.* 18, 1300–1307. <https://doi.org/10.1016/j.cub.2008.07.094>

828 Caro, T., 2014. Antipredator deception in terrestrial vertebrates. *Curr. Zool.* 60, 16–25.
829 <https://doi.org/10.1093/czoolo/60.1.16>

830 Cassey, T.C., Evens, D.R., Bogacz, R., Marshall, J.A.R., Ludwig, C.J.H., 2013. Adaptive Sampling of
831 Information in Perceptual Decision-Making. *PLOS ONE* 8, e78993.
832 <https://doi.org/10.1371/journal.pone.0078993>

833 Catania, K.C., 2011. The brain and behavior of the tentacled snake. *Ann. N. Y. Acad. Sci.* 1225, 83–89.
834 <https://doi.org/10.1111/j.1749-6632.2011.05959.x>

835 Chiovetto, E., Salatiello, A., d'Avella, A., Giese, M.A., 2022. Toward a unifying framework for the
836 modeling and identification of motor primitives. *Front. Comput. Neurosci.* 16, 926345.
837 <https://doi.org/10.3389/fncom.2022.926345>

838 Chittka, L., Skorupski, P., Raine, N.E., 2009. Speed–accuracy tradeoffs in animal decision making.
839 *Trends Ecol. Evol.* 24, 400–407. <https://doi.org/10.1016/j.tree.2009.02.010>

840 Churchland, M.M., Shenoy, K.V., 2024. Preparatory activity and the expansive null-space. *Nat. Rev.*
841 *Neurosci.* 25, 213–236. <https://doi.org/10.1038/s41583-024-00796-z>

842 Cisek, P., 2007. Cortical mechanisms of action selection: the affordance competition hypothesis.
843 *Philos. Trans. R. Soc. B Biol. Sci.* 362, 1585–1599. <https://doi.org/10.1098/rstb.2007.2054>

844 Cisek, P., Pastor-Bernier, A., 2014. On the challenges and mechanisms of embodied decisions. *Philos.*
845 *Trans. R. Soc. B Biol. Sci.* 369, 20130479. <https://doi.org/10.1098/rstb.2013.0479>

846 Claudi, F., Campagner, D., Branco, T., 2022. Innate heuristics and fast learning support escape route
847 selection in mice. *Curr. Biol.* 32, 2980-2987.e5. <https://doi.org/10.1016/j.cub.2022.05.020>

848 Collins, A.G.E., Cockburn, J., 2020. Beyond dichotomies in reinforcement learning. *Nat. Rev. Neurosci.*
849 21, 576–586. <https://doi.org/10.1038/s41583-020-0355-6>

850 Cooper Jr, W.E., 2006. Dynamic Risk Assessment: Prey Rapidly Adjust Flight Initiation Distance to
851 Changes in Predator Approach Speed. *Ethology* 112, 858–864. [https://doi.org/10.1111/j.1439-](https://doi.org/10.1111/j.1439-0310.2006.01240.x)
852 0310.2006.01240.x

853 Cooper Jr., W.E., 1997. Escape by a refuging prey, the broad-headed skink (*Eumeces laticeps*). *Can. J.*
854 *Zool.* 75, 943–947. <https://doi.org/10.1139/z97-113>

855 Cooper, W.E., Jr., 1998. Conditions favoring anticipatory and reactive displays deflecting predatory
856 attack. *Behav. Ecol.* 9, 598–604. <https://doi.org/10.1093/beheco/9.6.598>

857 Cos, I., Bélanger, N., Cisek, P., 2011. The influence of predicted arm biomechanics on decision making.
858 *J. Neurophysiol.* 105, 3022–3033. <https://doi.org/10.1152/jn.00975.2010>

859 Coulom, R., 2007. Efficient Selectivity and Backup Operators in Monte-Carlo Tree Search, in: Van Den
860 Herik, H.J., Ciancarini, P., Donkers, H.H.L.M. (Eds.), *Computers and Games, Lecture Notes in Computer*
861 *Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 72–83. [https://doi.org/10.1007/978-3-](https://doi.org/10.1007/978-3-540-75538-8_7)
862 540-75538-8_7

863 Courter, J.R., Ritchison, G., 2012. Asymmetries in Mobbing Behavior Among Nuclear Flockmates.
864 *Wilson J. Ornithol.* 124, 626–629. <https://doi.org/10.1676/11-168.1>

865 Davis, M., 2006. Neural systems involved in fear and anxiety measured with fear-potentiated startle.
866 *Am. Psychol.* 61, 741–756. <https://doi.org/10.1037/0003-066X.61.8.741>

867 Dayan, P., 2012. Instrumental vigour in punishment and reward. *Eur. J. Neurosci.* <https://doi.org/10.1111/J.1460-9568.2012.08026.X>

869 Devaine, M., Hollard, G., Daunizeau, J., 2014. Theory of Mind: Did Evolution Fool Us? *PLOS ONE* 9,
870 e87619. <https://doi.org/10.1371/journal.pone.0087619>

871 Dickinson, A., Balleine, B., 1994. Motivational Control of Goal-Directed Action. *Anim. Learn. Behav.*
872 <https://doi.org/10.3758/Bf03199951>

873 Dolan, R.J., Dayan, P., 2013. Goals and habits in the brain. *Neuron.*
874 <https://doi.org/10.1016/j.neuron.2013.09.007>

875 Domenici, P., 2010. Context-dependent variability in the components of fish escape response:
876 integrating locomotor performance and behavior. *J. Exp. Zool. Part Ecol. Genet. Physiol.* 313A, 59–79.
877 <https://doi.org/10.1002/jez.580>

878 Domenici, P., Blagburn, J.M., Bacon, J.P., 2011a. Animal escapology I: theoretical issues and emerging
879 trends in escape trajectories. *J. Exp. Biol.* 214, 2463–2473. <https://doi.org/10.1242/jeb.029652>

880 Domenici, P., Blagburn, J.M., Bacon, J.P., 2011b. Animal escapology II: escape trajectory case studies.
881 *J. Exp. Biol.* 214, 2474–2494. <https://doi.org/10.1242/jeb.053801>

882 Domenici, P., Booth, D., Blagburn, J.M., Bacon, J.P., 2008. Cockroaches Keep Predators Guessing by
883 Using Preferred Escape Trajectories. *Curr. Biol.* 18, 1792–1796.
884 <https://doi.org/10.1016/j.cub.2008.09.062>

885 Domenici, P., Hale, M.E., 2019. Escape responses of fish: a review of the diversity in motor control,
886 kinematics and behaviour. *J. Exp. Biol.* 222, jeb166009. <https://doi.org/10.1242/jeb.166009>

887 Duistermars, B.J., Pfeiffer, B.D., Hoopfer, E.D., Anderson, D.J., 2018. A Brain Module for Scalable
888 Control of Complex, Multi-motor Threat Displays. *Neuron* 100, 1474-1490.e4.
889 <https://doi.org/10.1016/j.neuron.2018.10.027>

890 Egner, T., 2009. Prefrontal cortex and cognitive control: motivating functional hierarchies. *Nat.*
891 *Neurosci.* 12, 821–822. <https://doi.org/10.1038/nn0709-821>

892 Eichenbaum, H., 2015. The Hippocampus as a Cognitive Map ... of Social Space. *Neuron* 87, 9–11.
893 <https://doi.org/10.1016/j.neuron.2015.06.013>

894 Enkhtaivan, E., Nishimura, J., Cochran, A., 2023. Placing Approach-Avoidance Conflict Within the
895 Framework of Multi-objective Reinforcement Learning. *Bull. Math. Biol.* 85, 116.
896 <https://doi.org/10.1007/s11538-023-01216-6>

897 Espinosa, G., Wink, G.E., Lai, A.T., Dombeck, D.A., MacIver, M.A., 2022. Achieving mouse-level strategic
898 evasion performance using real-time computational planning.
899 <https://doi.org/10.48550/arXiv.2211.02700>

900 Evans, D.A., Stempel, A.V., Vale, R., Branco, T., 2019. Cognitive Control of Escape Behaviour. *Trends*
901 *Cogn. Sci.* 23, 334–348. <https://doi.org/10.1016/j.tics.2019.01.012>

902 Evans, D.A., Stempel, A.V., Vale, R., Ruehle, S., Lefler, Y., Branco, T., 2018. A synaptic threshold
903 mechanism for computing escape decisions. *Nature*. <https://doi.org/10.1038/s41586-018-0244-6>

904 Fadok, J.P., Krabbe, S., Markovic, M., Courtin, J., Xu, C., Massi, L., Botta, P., Bylund, K., Muller, C.,
905 Kovacevic, A., Tovote, P., Luthi, A., 2017. A competitive inhibitory circuit for selection of active and
906 passive fear responses. *Nature*. <https://doi.org/10.1038/nature21047>

907 Fagen, R.M., 1976. Exercise, Play, and Physical Training in Animals, in: Bateson, P.P.G., Klopfer, P.H.
908 (Eds.), *Perspectives in Ethology: Volume 2*. Springer US, Boston, MA, pp. 189–219.
909 https://doi.org/10.1007/978-1-4615-7572-6_5

910 Fanselow, M.S., 1994. Neural organization of the defensive behavior system responsible for fear.
911 *Psychon. Bull. Rev.* 1, 429–438. <https://doi.org/10.3758/BF03210947>

912 Fanselow, M.S., Lester, L.S., 1988. A functional behavioristic approach to aversively motivated
913 behavior: Predatory imminence as a determinant of the topography of defensive behavior, in:
914 *Evolution and Learning*. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US, pp. 185–212.

915 Feher da Silva, C., Hare, T.A., 2020. Humans primarily use model-based inference in the two-stage
916 task. *Nat. Hum. Behav.* 4, 1053–1066. <https://doi.org/10.1038/s41562-020-0905-y>

917 Ferbinteanu, J., Shapiro, M.L., 2003. Prospective and Retrospective Memory Coding in the
918 Hippocampus. *Neuron* 40, 1227–1239. [https://doi.org/10.1016/S0896-6273\(03\)00752-9](https://doi.org/10.1016/S0896-6273(03)00752-9)

919 Fox, M., Gerevini, A., Long, D., Serina, I., 2006. Plan stability: replanning versus plan repair, in:
920 *Proceedings of the Sixteenth International Conference on International Conference on Automated*
921 *Planning and Scheduling, ICAPS'06*. AAAI Press, Cumbria, UK, pp. 212–221.

922 Fratzl, A., Koltchev, A.M., Vissers, N., Tan, Y.L., Marques-Smith, A., Stempel, A.V., Branco, T., Hofer,
923 S.B., 2021. Flexible inhibitory control of visually evoked defensive behavior by the ventral lateral
924 geniculate nucleus. *Neuron* 109, 3810–3822.e9. <https://doi.org/10.1016/j.neuron.2021.09.003>

925 Friston, K., 2010. The free-energy principle: a unified brain theory? *Nat Rev Neurosci.*
926 <https://doi.org/10.1038/nrn2787>

927 Frömer, R., Lin, H., Dean Wolf, C.K., Inzlicht, M., Shenhav, A., 2021. Expectations of reward and efficacy
928 guide cognitive control allocation. *Nat. Commun.* 12, 1030. [https://doi.org/10.1038/s41467-021-](https://doi.org/10.1038/s41467-021-21315-z)
929 [21315-z](https://doi.org/10.1038/s41467-021-21315-z)

930 Gagne, C., Dayan, P., 2022. Peril, prudence and planning as risk, avoidance and worry. *J. Math. Psychol.*
931 106, 102617. <https://doi.org/10.1016/j.jmp.2021.102617>

932 Gallivan, J.P., Barton, K.S., Chapman, C.S., Wolpert, D.M., Randall Flanagan, J., 2015. Action plan co-
933 optimization reveals the parallel encoding of competing reach movements. *Nat. Commun.* 6, 7428.
934 <https://doi.org/10.1038/ncomms8428>

935 Gallivan, J.P., Chapman, C.S., Wolpert, D.M., Flanagan, J.R., 2018. Decision-making in sensorimotor
936 control. *Nat. Rev. Neurosci.* 19, 519–534. <https://doi.org/10.1038/s41583-018-0045-9>

937 Gershman, S.J., 2020. Origin of perseveration in the trade-off between reward and complexity.
938 *Cognition* 204, 104394. <https://doi.org/10.1016/j.cognition.2020.104394>

939 Gibson, W.T., Gonzalez, C.R., Fernandez, C., Ramasamy, L., Tabachnik, T., Du, R.R., Felsen, P.D., Maire,
940 M.R., Perona, P., Anderson, D.J., 2015. Behavioral Responses to a Repetitive Visual Threat Stimulus
941 Express a Persistent State of Defensive Arousal in *Drosophila*. *Curr. Biol.* 25, 1401–1415.
942 <https://doi.org/10.1016/j.cub.2015.03.058>

943 Gilbert, C., 1997. Visual control of cursorial prey pursuit by tiger beetles (Cicindelidae). *J. Comp.*
944 *Physiol. A* 181, 217–230. <https://doi.org/10.1007/s003590050108>

945 Gmytrasiewicz, P.J., Doshi, P., 2005. A framework for sequential planning in multi-agent settings. *J.*
946 *Artif. Intell. Res.* 24, 49–79.

947 Gold, J.I., Stocker, A.A., 2017. Visual Decision-Making in an Uncertain and Dynamic World. *Annu. Rev.*
948 *Vis. Sci.* 3, 227–250. <https://doi.org/10.1146/annurev-vision-111815-114511>

949 Gordon, J., Maselli, A., Lancia, G.L., Thiery, T., Cisek, P., Pezzulo, G., 2021. The road towards
950 understanding embodied decisions. *Neurosci. Biobehav. Rev.* 131, 722–736.
951 <https://doi.org/10.1016/j.neubiorev.2021.09.034>

952 Gottlieb, J., Oudeyer, P.-Y., Lopes, M., Baranes, A., 2013. Information-seeking, curiosity, and attention:
953 computational and neural mechanisms. *Trends Cogn. Sci.* 17, 585–593.
954 <https://doi.org/10.1016/j.tics.2013.09.001>

955 Griffin, A.S., Evans, C.S., Blumstein, D.T., 2001. Learning specificity in acquired predator recognition.
956 *Anim. Behav.* 62, 577–589. <https://doi.org/10.1006/anbe.2001.1781>

957 Guinet, C., Domenici, P., De Stephanis, R., Barrett-Lennard, L., Ford, J., Verborgh, P., 2007. Killer whale
958 predation on bluefin tuna: exploring the hypothesis of the endurance-exhaustion technique. *Mar.*
959 *Ecol. Prog. Ser.* 347, 111–119. <https://doi.org/10.3354/meps07035>

960 Guo, J., Xia, W., Hu, X., Ma, H., 2022. Feedback RRT* algorithm for UAV path planning in a hostile
961 environment. *Comput. Ind. Eng.* 174, 108771. <https://doi.org/10.1016/j.cie.2022.108771>

962 Guzman, C., Castejon, P., Onaindia, E., Frank, J., 2014. Robust Plan Execution in Multi-agent
963 Environments, in: 2014 IEEE 26th International Conference on Tools with Artificial Intelligence.
964 Presented at the 2014 IEEE 26th International Conference on Tools with Artificial Intelligence (ICTAI),
965 IEEE, Limassol, Cyprus, pp. 384–391. <https://doi.org/10.1109/ICTAI.2014.65>

966 Hassenstein, B., Hustert, R., 1999. Hiding responses of locusts to approaching objects. *J. Exp. Biol.* 202,
 967 1701–1710. <https://doi.org/10.1242/jeb.202.12.1701>

968 Hasson, C.J., Shen, T., Sternad, D., 2012. Energy margins in dynamic object manipulation. *J.*
 969 *Neurophysiol.* 108, 1349–1365. <https://doi.org/10.1152/jn.00019.2012>

970 Hausken, K., Levitin, G., 2009. Minmax defense strategy for complex multi-state systems. *Reliab. Eng.*
 971 *Syst. Saf.* 94, 577–587. <https://doi.org/10.1016/j.res.2008.06.005>

972 Hay, N., Russell, S., Tolpin, D., Shimony, S.E., 2012. Selecting computations: theory and applications,
 973 in: *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence, UAI'12*. AUAI
 974 Press, Arlington, Virginia, USA, pp. 346–355.

975 Hein, A.M., Altshuler, D.L., Cade, D.E., Liao, J.C., Martin, B.T., Taylor, G.K., 2020. An Algorithmic
 976 Approach to Natural Behavior. *Curr. Biol.* 30, R663–R675. <https://doi.org/10.1016/j.cub.2020.04.018>

977 Hemmi, J.M., 2005. Predator avoidance in fiddler crabs: 2. The visual cues. *Anim. Behav.* 69, 615–625.
 978 <https://doi.org/10.1016/j.anbehav.2004.06.019>

979 Hemmi, J.M., Pfeil, A., 2010. A multi-stage anti-predator response increases information on predation
 980 risk. *J. Exp. Biol.* 213, 1484–1489. <https://doi.org/10.1242/jeb.039925>

981 Heng, J.A., Woodford, M., Polania, R., 2020. Efficient sampling and noisy decisions. *eLife* 9, e54962.
 982 <https://doi.org/10.7554/eLife.54962>

983 Herbert-Read, J.E., Ward, A.J.W., Sumpter, D.J.T., Mann, R.P., 2017. Escape path complexity and its
 984 context dependency in Pacific blue-eyes (*Pseudomugil signifer*). *J. Exp. Biol.* jeb.154534.
 985 <https://doi.org/10.1242/jeb.154534>

986 Herry, C., Ciocchi, S., Senn, V., Demmou, L., Müller, C., Lüthi, A., 2008. Switching on and off fear by
 987 distinct neuronal circuits. *Nature* 454, 600–606. <https://doi.org/10.1038/nature07166>

988 Howland, H.C., 1974. Optimal strategies for predator avoidance: The relative importance of speed and
 989 manoeuvrability. *J. Theor. Biol.* 47, 333–350. [https://doi.org/10.1016/0022-5193\(74\)90202-1](https://doi.org/10.1016/0022-5193(74)90202-1)

990 Humphries, D.A., Driver, P.M., 1970. Protean defence by prey animals. *Oecologia* 5, 285–302.
 991 <https://doi.org/10.1007/BF00815496>

992 Huys, Q.J., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S.J., Dayan, P., Roiser, J.P., 2015.
 993 Interplay of approximate planning strategies. *Proc Natl Acad Sci U A.*
 994 <https://doi.org/10.1073/pnas.1414219112>

995 Huys, Q.J.M., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S.J., Dayan, P., Roiser, J.P., 2015.
 996 Interplay of approximate planning strategies. *Proc. Natl. Acad. Sci.* 112, 3098–3103.
 997 <https://doi.org/10.1073/pnas.1414219112>

998 Janssen, M., LeWarne, C., Burk, D., Averbeck, B.B., 2022. Hierarchical Reinforcement Learning,
 999 Sequential Behavior, and the Dorsal Frontostriatal System. *J. Cogn. Neurosci.* 34, 1307–1325.
 1000 https://doi.org/10.1162/jocn_a_01869

1001 Jornod, M., Roche, D.G., 2015. Inter- vs intra-individual variation and temporal repeatability of escape
 1002 responses in the coral reef fish *Amblyglyphidodon curacao*. *Biol. Open* 4, 1395–1399.
 1003 <https://doi.org/10.1242/bio.013508>

1004 Kavaliers, M., Choleris, E., 2001. Antipredator responses and defensive behavior: ecological and
 1005 ethological approaches for the neurosciences. *Neurosci. Biobehav. Rev.* 25, 577–586.
 1006 [https://doi.org/10.1016/S0149-7634\(01\)00042-2](https://doi.org/10.1016/S0149-7634(01)00042-2)

1007 Kawabata, Y., Akada, H., Shimatani, K., Nishihara, G.N., Kimura, H., Nishiumi, N., Domenici, P., 2023.
 1008 Multiple preferred escape trajectories are explained by a geometric model incorporating prey's turn
 1009 and predator attack endpoint. *eLife* 12, e77699. <https://doi.org/10.7554/eLife.77699>

1010 Kearns, M., Mansour, Y., Ng, A., 1999. Approximate planning in large POMDPs via reusable
 1011 trajectories. *Adv. Neural Inf. Process. Syst.* 12.

1012 Kearns, M., Mansour, Y., Ng, A.Y., 2002. A Sparse Sampling Algorithm for Near-Optimal Planning in
 1013 Large Markov Decision Processes. *Mach. Learn.* 49, 193–208.
 1014 <https://doi.org/10.1023/A:1017932429737>

1015 Kennedy, A., Kunwar, P.S., Li, L., Stagkourakis, S., Wagenaar, D.A., Anderson, D.J., 2020. Stimulus-
 1016 specific hypothalamic encoding of a persistent defensive state. *Nature* 586, 730–734.
 1017 <https://doi.org/10.1038/s41586-020-2728-4>

1018 Keramati, M., Gutkin, B., 2014. Homeostatic reinforcement learning for integrating reward collection
 1019 and physiological stability. *Elife*. <https://doi.org/10.7554/eLife.04811>

1020 Keramati, M., Gutkin, B., 2011. A Reinforcement Learning Theory for Homeostatic Regulation, in:
 1021 Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., Weinberger, K.Q. (Eds.), *Advances in Neural*
 1022 *Information Processing Systems*. Curran Associates, Inc.

1023 Keramati, M., Smittenaar, P., Dolan, R.J., Dayan, P., 2016. Adaptive integration of habits into depth-
 1024 limited planning defines a habitual-goal-directed spectrum. *Proc. Natl. Acad. Sci.* 113, 12868–12873.
 1025 <https://doi.org/10.1073/pnas.1609094113>

1026 Khatib, O., 1987. A unified approach for motion and force control of robot manipulators: The
 1027 operational space formulation. IEEE J. Robot. Autom. 3, 43–53.
 1028 <https://doi.org/10.1109/JRA.1987.1087068>

1029 Kimura, H., Pfalzgraff, T., Levet, M., Kawabata, Y., Steffensen, J.F., Johansen, J.L., Domenici, P., 2022.
 1030 Escaping from multiple visual threats: modulation of escape responses in Pacific staghorn sculpin (
 1031 *Leptocottus armatus*). J. Exp. Biol. 225, jeb243328. <https://doi.org/10.1242/jeb.243328>

1032 Klioutchnikov, A., Wallace, D.J., Sawinski, J., Voit, K.-M., Groemping, Y., Kerr, J.N.D., 2023. A three-
 1033 photon head-mounted microscope for imaging all layers of visual cortex in freely moving mice. Nat.
 1034 Methods 20, 610–616. <https://doi.org/10.1038/s41592-022-01688-9>

1035 Koechlin, E., Hyafil, A., 2007. Anterior Prefrontal Function and the Limits of Human Decision-Making.
 1036 Science 318, 594–598. <https://doi.org/10.1126/science.1142995>

1037 Kopetz, H., Steiner, W., 2022. Real-time systems: design principles for distributed embedded
 1038 applications. Springer Nature.

1039 Lai, L., Gershman, S.J., 2021. Policy compression: An information bottleneck in action selection, in:
 1040 Psychology of Learning and Motivation. Elsevier, pp. 195–232.
 1041 <https://doi.org/10.1016/bs.plm.2021.02.004>

1042 Laing, P.A.F., Vervliet, B., Dunsmoor, J.E., Harrison, B.J., 2024. Pavlovian safety learning: An integrative
 1043 theoretical review. Psychon. Bull. Rev. <https://doi.org/10.3758/s13423-024-02559-4>

1044 Latash, M.L., 2021. Laws of nature that define biological action and perception. Phys. Life Rev. 36, 47–
 1045 67. <https://doi.org/10.1016/j.plrev.2020.07.007>

1046 LeDoux, J., Daw, N.D., 2018. Surviving threats: neural circuit and computational implications of a new
 1047 taxonomy of defensive behaviour. Nat. Rev. Neurosci. 19, 269–282.
 1048 <https://doi.org/10.1038/nrn.2018.22>

1049 LeDoux, J.E., 1995. Emotion: clues from the brain. Annu. Rev. Psychol.

1050 Lee, S., Hwang, S., Joe, Y., Cha, H., Joo, G., Lee, H., Kim, J., Jablonski, P.G., 2013. Direct Look from a
 1051 Predator Shortens the Risk-Assessment Time by Prey. PLoS ONE 8, e64977.
 1052 <https://doi.org/10.1371/journal.pone.0064977>

1053 Li, W., 2017. A Dynamics Perspective of Pursuit-Evasion: Capturing and Escaping When the Pursuer
 1054 Runs Faster Than the Agile Evader. IEEE Trans. Autom. Control 62, 451–457.
 1055 <https://doi.org/10.1109/TAC.2016.2575008>

1056 Liden, W.H., Phillips, M.L., Herberholz, J., 2010. Neural control of behavioural choice in juvenile
 1057 crayfish. *Proc. R. Soc. B Biol. Sci.* 277, 3493–3500. <https://doi.org/10.1098/rspb.2010.1000>
 1058 Lieder, F., Griffiths, T.L., 2020. Resource-rational analysis: Understanding human cognition as the
 1059 optimal use of limited computational resources. *Behav. Brain Sci.* 43, e1.
 1060 <https://doi.org/10.1017/S0140525X1900061X>
 1061 Lieder, F., Hsu, M., Griffiths, T.L., 2014. The high availability of extreme events serves resource-rational
 1062 decision-making. *Proc. Annu. Meet. Cogn. Sci. Soc.* 36.
 1063 Lieder, F., Shenhav, A., Musslick, S., Griffiths, T.L., 2018. Rational metareasoning and the plasticity of
 1064 cognitive control. *PLOS Comput. Biol.* 14, e1006043. <https://doi.org/10.1371/journal.pcbi.1006043>
 1065 Lima, S.L., Dill, L.M., 1990. Behavioral decisions made under the risk of predation: a review and
 1066 prospectus. *Can. J. Zool.*
 1067 Liu, Y., Peng, Y., Wang, B., Yao, S., Liu, Z., 2017. Review on cyber-physical systems. *IEEECAA J. Autom.*
 1068 *Sin.* 4, 27–40. <https://doi.org/10.1109/JAS.2017.7510349>
 1069 LoBue, V., Matthews, K., Harvey, T., Stark, S.L., 2014. What accounts for the rapid detection of threat?
 1070 Evidence for an advantage in perceptual and behavioral responding from eye movements. *Emotion*
 1071 14, 816–823. <https://doi.org/10.1037/a0035869>
 1072 Lu, C., Schölkopf, B., Hernández-Lobato, J.M., 2018. Deconfounding Reinforcement Learning in
 1073 Observational Settings. <https://doi.org/10.48550/ARXIV.1812.10576>
 1074 Luo, Y., Bai, H., Hsu, D., Lee, W.S., 2019. Importance sampling for online planning under uncertainty.
 1075 *Int. J. Robot. Res.* 38, 162–181. <https://doi.org/10.1177/0278364918780322>
 1076 Majumdar, A., Tedrake, R., 2013. Robust Online Motion Planning with Regions of Finite Time
 1077 Invariance, in: Frazzoli, E., Lozano-Perez, T., Roy, N., Rus, D. (Eds.), *Algorithmic Foundations of Robotics*
 1078 X, Springer Tracts in Advanced Robotics. Springer, Berlin, Heidelberg, pp. 543–558.
 1079 https://doi.org/10.1007/978-3-642-36279-8_33
 1080 Mandralis, I., Weber, P., Novati, G., Koumoutsakos, P., 2021. Learning swimming escape patterns for
 1081 larval fish under energy constraints. *Phys. Rev. Fluids* 6, 093101.
 1082 <https://doi.org/10.1103/PhysRevFluids.6.093101>
 1083 Manohar, S.G., Chong, T.T.-J., Apps, M.A.J., Batla, A., Stamelou, M., Jarman, P.R., Bhatia, K.P., Husain,
 1084 M., 2015. Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. *Curr. Biol.* 25,
 1085 1707–1716. <https://doi.org/10.1016/j.cub.2015.05.038>

1086 Marois, R., Ivanoff, J., 2005. Capacity limits of information processing in the brain. *Trends Cogn. Sci.* 9,
1087 296–305. <https://doi.org/10.1016/j.tics.2005.04.010>

1088 Marr, D.C., Poggio, T., 1977. From Understanding Computation to Understanding Neural Circuitry.
1089 *Neurosci. Res. Program Bull.*

1090 Marras, S., Domenici, P., 2013. Schooling Fish Under Attack Are Not All Equal: Some Lead, Others
1091 Follow. *PLOS ONE* 8, e65784. <https://doi.org/10.1371/journal.pone.0065784>

1092 Mattar, M.G., Daw, N.D., 2018. Prioritized memory access explains planning and hippocampal replay.
1093 *Nat. Neurosci.* 21, 1609–1617. <https://doi.org/10.1038/s41593-018-0232-z>

1094 Mattar, M.G., Lengyel, M., 2022. Planning in the brain. *Neuron* 110, 914–934.
1095 <https://doi.org/10.1016/j.neuron.2021.12.018>

1096 Merel, J., Botvinick, M., Wayne, G., 2019. Hierarchical motor control in mammals and machines. *Nat.*
1097 *Commun.* 10, 5489. <https://doi.org/10.1038/s41467-019-13239-6>

1098 Mern, J., Yildiz, A., Bush, L., Mukerji, T., Kochenderfer, M.J., 2021. Improved POMDP Tree Search
1099 Planning with Prioritized Action Branching. *Proc. AAAI Conf. Artif. Intell.* 35, 11888–11894.
1100 <https://doi.org/10.1609/aaai.v35i13.17412>

1101 Mobbs, D., Headley, D.B., Ding, W., Dayan, P., 2020. Space, Time, and Fear: Survival Computations
1102 along Defensive Circuits. *Trends Cogn. Sci.* 24, 228–241. <https://doi.org/10.1016/j.tics.2019.12.016>

1103 Mobbs, D., Marchant, J.L., Hassabis, D., Seymour, B., Tan, G., Gray, M., Petrovic, P., Dolan, R.J., Frith,
1104 C.D., 2009. From Threat to Fear: The Neural Organization of Defensive Fear Systems in Humans. *J.*
1105 *Neurosci.* 29, 12236–12243. <https://doi.org/10.1523/JNEUROSCI.2378-09.2009>

1106 Mohammad, F., Aryal, S., Ho, J., Stewart, J.C., Norman, N.A., Tan, T.L., Eisaka, A., Claridge-Chang, A.,
1107 2016. Ancient Anxiety Pathways Influence *Drosophila* Defense Behaviors. *Curr. Biol.* 26, 981–986.
1108 <https://doi.org/10.1016/j.cub.2016.02.031>

1109 Musslick, S., Cohen, J.D., 2021. Rationalizing constraints on the capacity for cognitive control. *Trends*
1110 *Cogn. Sci.* 25, 757–775. <https://doi.org/10.1016/j.tics.2021.06.001>

1111 Nagengast, A.J., Braun, D.A., Wolpert, D.M., 2010. Risk-Sensitive Optimal Feedback Control Accounts
1112 for Sensorimotor Behavior under Uncertainty. *PLOS Comput. Biol.* 6, e1000857.
1113 <https://doi.org/10.1371/journal.pcbi.1000857>

1114 Niu, S., Liu, Y., Wang, J., Song, H., 2020. A Decade Survey of Transfer Learning (2010–2020). *IEEE Trans.*
1115 *Artif. Intell.* 1, 151–166. <https://doi.org/10.1109/TAI.2021.3054609>

1116 Nord, C.L., Prabhu, G., Nolte, T., Fonagy, P., Dolan, R., Moutoussis, M., 2017. Vigour in active
 1117 avoidance. *Sci. Rep.* 7, 60. <https://doi.org/10.1038/s41598-017-00127-6>
 1118 O'Keefe, J.M., Nadel, L., O'Keefe, J., 1978. The hippocampus as a cognitive map. Clarendon Press,
 1119 Oxford.
 1120 Olafsdottir, H.F., Bush, D., Barry, C., 2018. The Role of Hippocampal Replay in Memory and Planning.
 1121 *Curr Biol.* <https://doi.org/10.1016/j.cub.2017.10.073>
 1122 Olsson, A., Knapska, E., Lindstrom, B., 2020. The neural and computational systems of social learning.
 1123 *Nat Rev Neurosci* 21, 197–212. <https://doi.org/10.1038/s41583-020-0276-4>
 1124 O'Neill, G.C., Seymour, R.A., Mellor, S., Alexander, N.A., Tierney, T.M., Bernachot, L., Fahimi Hnazaee,
 1125 M., Spedden, M.E., Timms, R.C., Bush, D., Bestmann, S., Brookes, M.J., Barnes, G.R., 2025. Combining
 1126 video telemetry and wearable MEG for naturalistic imaging. *Imaging Neurosci.* 3, imag_a_00495.
 1127 https://doi.org/10.1162/imag_a_00495
 1128 Ott, F., Marković, D., Strobel, A., Kiebel, S.J., 2020. Dynamic integration of forward planning and
 1129 heuristic preferences during multiple goal pursuit. *PLOS Comput. Biol.* 16, e1007685.
 1130 <https://doi.org/10.1371/journal.pcbi.1007685>
 1131 Ozbagci, D., Moreno-Bote, R., Soto-Faraco, S., 2021. The dynamics of decision-making and action
 1132 during active sampling. *Sci. Rep.* 11, 23067. <https://doi.org/10.1038/s41598-021-02595-3>
 1133 Panoz-Brown, D., Iyer, V., Carey, L.M., Sluka, C.M., Rajic, G., Kestenman, J., Gentry, M., Brotheridge,
 1134 S., Somekh, I., Corbin, H.E., Tucker, K.G., Almeida, B., Hex, S.B., Garcia, K.D., Hohmann, A.G., Crystal,
 1135 J.D., 2018. Replay of Episodic Memories in the Rat. *Curr. Biol.* CB 28, 1628-1634.e7.
 1136 <https://doi.org/10.1016/j.cub.2018.04.006>
 1137 Peterson, A.N., Soto, A.P., McHenry, M.J., 2021. Pursuit and Evasion Strategies in the Predator–Prey
 1138 Interactions of Fishes. *Integr. Comp. Biol.* 61, 668–680. <https://doi.org/10.1093/icb/icab116>
 1139 Petri, G., Musslick, S., Dey, B., Özcimder, K., Turner, D., Ahmed, N.K., Willke, T.L., Cohen, J.D., 2021.
 1140 Topological limits to the parallel processing capability of network architectures. *Nat. Phys.* 17, 646–
 1141 651. <https://doi.org/10.1038/s41567-021-01170-x>
 1142 Pezzulo, G., Cisek, P., 2016. Navigating the Affordance Landscape: Feedback Control as a Process
 1143 Model of Behavior and Cognition. *Trends Cogn Sci.* <https://doi.org/10.1016/j.tics.2016.03.013>
 1144 Pfeiffer, B.E., Foster, D.J., 2013. Hippocampal place-cell sequences depict future paths to remembered
 1145 goals. *Nature* 497, 74–79. <https://doi.org/10.1038/nature12112>

1146 Piray, P., Daw, N.D., 2021. Linear reinforcement learning in planning, grid fields, and cognitive control.
 1147 Nat. Commun. 12, 4942. <https://doi.org/10.1038/s41467-021-25123-3>

1148 Prystawski, B., Mohnert, F., Tošić, M., Lieder, F., 2022. Resource-rational Models of Human Goal
 1149 Pursuit. Top. Cogn. Sci. 14, 528–549. <https://doi.org/10.1111/tops.12562>

1150 Puelma Touzel, M., Cisek, P., Lajoie, G., 2022. Performance-gated deliberation: A context-adapted
 1151 strategy in which urgency is opportunity cost. PLOS Comput. Biol. 18, e1010080.
 1152 <https://doi.org/10.1371/journal.pcbi.1010080>

1153 Rao, R.P.N., 2010. Decision Making Under Uncertainty: A Neural Model Based on Partially Observable
 1154 Markov Decision Processes. Front. Comput. Neurosci. 4. <https://doi.org/10.3389/fncom.2010.00146>

1155 Ratcliff, R., Smith, P.L., Brown, S.D., McKoon, G., 2016. Diffusion Decision Model: Current Issues and
 1156 History. Trends Cogn. Sci. 20, 260–281. <https://doi.org/10.1016/j.tics.2016.01.007>

1157 Reynaud, A.J., Saleri Lunazzi, C., Thura, D., 2020. Humans sacrifice decision-making for action
 1158 execution when a demanding control of movement is required. J. Neurophysiol. 124, 497–509.
 1159 <https://doi.org/10.1152/jn.00220.2020>

1160 Roberts, G., 1996. Why individual vigilance declines as group size increases. Anim. Behav. 51, 1077–
 1161 1086. <https://doi.org/10.1006/anbe.1996.0109>

1162 Roberts, G., Holmes, N., Alexander, N., Boto, E., Leggett, J., Hill, R.M., Shah, V., Rea, M., Vaughan, R.,
 1163 Maguire, E.A., Kessler, K., Beebe, S., Fromhold, M., Barnes, G.R., Bowtell, R., Brookes, M.J., 2019.
 1164 Towards OPM-MEG in a virtual reality environment. NeuroImage 199, 408–417.
 1165 <https://doi.org/10.1016/j.neuroimage.2019.06.010>

1166 Saitoh, K., Ménard, A., Grillner, S., 2007. Tectal Control of Locomotion, Steering, and Eye Movements
 1167 in Lamprey. J. Neurophysiol. 97, 3093–3108. <https://doi.org/10.1152/jn.00639.2006>

1168 Salvanes, A.G.V., Moberg, O., Ebbesson, L.O.E., Nilsen, T.O., Jensen, K.H., Braithwaite, V.A., 2013.
 1169 Environmental enrichment promotes neural plasticity and cognitive ability in fish. Proc. R. Soc. B Biol.
 1170 Sci. 280, 20131331. <https://doi.org/10.1098/rspb.2013.1331>

1171 Santer, R.D., Yamawaki, Y., Rind, F.C., Simmons, P.J., 2005. Motor activity and trajectory control during
 1172 escape jumping in the locust *Locusta migratoria*. J. Comp. Physiol. A 191, 965–975.
 1173 <https://doi.org/10.1007/s00359-005-0023-3>

1174 Sato, N., Shidara, H., Ogawa, H., 2019. Trade-off between motor performance and behavioural
 1175 flexibility in the action selection of cricket escape behaviour. Sci. Rep. 9, 18112.
 1176 <https://doi.org/10.1038/s41598-019-54555-7>

1177 Savage, L.J., 1954. The Foundations of Statistics.

1178 Scala, E., Micalizio, R., Torasso, P., 2015. Robust plan execution via reconfiguration
1179 and replanning. *AI Commun.* 28, 479–509. <https://doi.org/10.3233/AIC-140629>

1180 Schmidt-Daffy, M., 2011. Modeling automatic threat detection: development of a face-in-the-crowd
1181 task. *Emotion*. <https://doi.org/10.1037/a0022018>

1182 Selen, L.P.J., Shadlen, M.N., Wolpert, D.M., 2012. Deliberation in the Motor System: Reflex Gains Track
1183 Evolving Evidence Leading to a Decision. *J. Neurosci.* 32, 2276–2286.
1184 <https://doi.org/10.1523/JNEUROSCI.5273-11.2012>

1185 Sennesh, E., Theriault, J., Brooks, D., Van De Meent, J.-W., Barrett, L.F., Quigley, K.S., 2022.
1186 Interoception as modeling, allostasis as control. *Biol. Psychol.* 167, 108242.
1187 <https://doi.org/10.1016/j.biopsycho.2021.108242>

1188 Shadlen, M.N., Kiani, R., 2013. Decision Making as a Window on Cognition. *Neuron* 80, 791–806.
1189 <https://doi.org/10.1016/j.neuron.2013.10.047>

1190 Shadmehr, R., 2020. Vigor: neuroeconomics of movement control. The MIT Press, Cambridge,
1191 Massachusetts.

1192 Shadmehr, R., Huang, H.J., Ahmed, A.A., 2016. A Representation of Effort in Decision-Making and
1193 Motor Control. *Curr. Biol.* 26, 1929–1934. <https://doi.org/10.1016/j.cub.2016.05.065>

1194 Shamash, P., Olesen, S.F., Iordanidou, P., Campagner, D., Banerjee, N., Branco, T., 2021. Mice learn
1195 multi-step routes by memorizing subgoal locations. *Nat. Neurosci.* 24, 1270–1279.
1196 <https://doi.org/10.1038/s41593-021-00884-8>

1197 Shea-Brown, E., Gilzenrat, M.S., Cohen, J.D., 2008. Optimization of Decision Making in Multilayer
1198 Networks: The Role of Locus Coeruleus. *Neural Comput.* 20, 2863–2894.
1199 <https://doi.org/10.1162/neco.2008.03-07-487>

1200 Shenhav, A., Botvinick, M.M., Cohen, J.D., 2013. The Expected Value of Control: An Integrative Theory
1201 of Anterior Cingulate Cortex Function. *Neuron* 79, 217–240.
1202 <https://doi.org/10.1016/j.neuron.2013.07.007>

1203 Shi, X., Møller, J.S., Højgaard, J., Johansen, J.L., Steffensen, J.F., Liu, D., Domenici, P., 2017. The angular
1204 position of a refuge affects escape responses in staghorn sculpin *Leptocottus armatus*. *J. Fish Biol.* 90,
1205 2434–2442. <https://doi.org/10.1111/jfb.13306>

1206 Sih, A., 1992. Prey Uncertainty and the Balancing of Antipredator and Feeding Needs. *Am. Nat.* 139,
1207 1052–1069. <https://doi.org/10.1086/285372>

1208 Simon, H.A., 1979. Rational decision making in business organizations. *Am. Econ. Rev.* 69, 493–513.

1209 Sinha, A., O’Kelly, M., Zheng, H., Mangharam, R., Duchi, J., Tedrake, R., 2020. FormulaZero:
 1210 Distributionally Robust Online Adaptation via Offline Population Synthesis, in: *Proceedings of the 37th*
 1211 *International Conference on Machine Learning*. Presented at the International Conference on
 1212 *Machine Learning*, PMLR, pp. 8992–9004.

1213 Smith, M.L., Asada, N., Malenka, R.C., 2021. Anterior cingulate inputs to nucleus accumbens control
 1214 the social transfer of pain and analgesia. *Science* 371, 153–159.
 1215 <https://doi.org/10.1126/science.abe3040>

1216 Smith, P.L., Saber, S., Corbett, E.A., Lilburn, S.D., 2020. Modeling continuous outcome color decisions
 1217 with the circular diffusion model: Metric and categorical properties. *Psychol. Rev.* 127, 562–590.
 1218 <https://doi.org/10.1037/rev0000185>

1219 Snider, J., Ahmed, O.J., Halgren, E., Poizner, H., Cash, S.S., 2013. Human intracranial recordings during
 1220 spatial exploration of a 3D virtual environment, in: *2013 6th International IEEE/EMBS Conference on*
 1221 *Neural Engineering (NER)*. Presented at the 2013 6th International IEEE/EMBS Conference on Neural
 1222 *Engineering (NER)*, pp. 464–467. <https://doi.org/10.1109/NER.2013.6695972>

1223 Snider, J., Lee, D., Poizner, H., Gepshtein, S., 2015. Prospective Optimization with Limited Resources.
 1224 *PLOS Comput. Biol.* 11, e1004501. <https://doi.org/10.1371/journal.pcbi.1004501>

1225 Soto, A., Stewart, W.J., McHenry, M.J., 2015. When Optimal Strategy Matters to Prey Fish. *Integr.*
 1226 *Comp. Biol.* 55, 110–120. <https://doi.org/10.1093/icb/icv027>

1227 Spinka, M., Newberry, R.C., Bekoff, M., 2001. Mammalian Play: Training for the Unexpected. *Q. Rev.*
 1228 *Biol.* 76, 141–168. <https://doi.org/10.1086/393866>

1229 Sporrer, J.K., Brookes, J., Hall, S., Zabbah, S., Serratos Hernandez, U.D., Bach, D.R., 2023. Functional
 1230 sophistication in human escape. *iScience* 26, 108240. <https://doi.org/10.1016/j.isci.2023.108240>

1231 Stempel, A.V., Evans, D.A., Arocas, O.P., Claudi, F., Lenzi, S.C., Kutsarova, E., Margrie, T.W., Branco, T.,
 1232 2024. Tonically active GABAergic neurons in the dorsal periaqueductal gray control instinctive escape
 1233 in mice. *Curr. Biol.* 34, 3031–3039.e7. <https://doi.org/10.1016/j.cub.2024.05.068>

1234 Stephens, D.W., 2008. Decision ecology: Foraging and the ecology of animal decision making. *Cogn.*
 1235 *Affect. Behav. Neurosci.* 8, 475–484. <https://doi.org/10.3758/CABN.8.4.475>

1236 Stone, C., Mattingley, J.B., Rangelov, D., 2022. On second thoughts: changes of mind in decision-
 1237 making. *Trends Cogn. Sci.* 26, 419–431. <https://doi.org/10.1016/j.tics.2022.02.004>

1238 Storms, R.F., Carere, C., Zoratto, F., Hemelrijk, C.K., 2019. Complex patterns of collective escape in
1239 starling flocks under predation. *Behav. Ecol. Sociobiol.* 73, 10. <https://doi.org/10.1007/s00265-018->
1240 2609-0

1241 Sutton, R.S., 1990. Integrated Architectures for Learning, Planning, and Reacting Based on
1242 Approximating Dynamic Programming, in: Porter, B., Mooney, R. (Eds.), *Machine Learning Proceedings*
1243 1990. Morgan Kaufmann, San Francisco (CA), pp. 216–224. <https://doi.org/10.1016/B978-1-55860->
1244 141-3.50030-4

1245 Tätte, K., Møller, A.P., Mänd, R., 2020. Corvids exhibit dynamic risk assessment during escape. *Behav.*
1246 *Processes* 170, 104017. <https://doi.org/10.1016/j.beproc.2019.104017>

1247 Taylor, A.H., Bastos, A.P.M., Brown, R.L., Allen, C., 2022. The signature-testing approach to mapping
1248 biological and artificial intelligences. *Trends Cogn. Sci.* 26, 738–750.
1249 <https://doi.org/10.1016/j.tics.2022.06.002>

1250 Ten, A., Kaushik, P., Oudeyer, P.-Y., Gottlieb, J., 2021. Humans monitor learning progress in curiosity-
1251 driven exploration. *Nat. Commun.* 12, 5972. <https://doi.org/10.1038/s41467-021-26196-w>

1252 Tetzlaff, S.J., Sperry, J.H., DeGregorio, B.A., 2019. Effects of antipredator training, environmental
1253 enrichment, and soft release on wildlife translocations: A review and meta-analysis. *Biol. Conserv.* 236,
1254 324–331. <https://doi.org/10.1016/j.biocon.2019.05.054>

1255 Thorndike, E.L., 1911. *Animal intelligence: Experimental studies*. Macmillan, New York.

1256 Thura, D., Beauregard-Racine, J., Fradet, C.-W., Cisek, P., 2012. Decision making by urgency gating:
1257 theory and experimental support. *J. Neurophysiol.* 108, 2912–2930.
1258 <https://doi.org/10.1152/jn.01071.2011>

1259 Thura, D., Cisek, P., 2017. The Basal Ganglia Do Not Select Reach Targets but Control the Urgency of
1260 Commitment. *Neuron* 95, 1160-1170.e5. <https://doi.org/10.1016/j.neuron.2017.07.039>

1261 Todorov, E., 2004. Optimality principles in sensorimotor control. *Nat. Neurosci.* 7, 907–915.
1262 <https://doi.org/10.1038/nn1309>

1263 Tolman, E.C., 1948. Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.
1264 <https://doi.org/10.1037/h0061626>

1265 Tombu, M.N., Asplund, C.L., Dux, P.E., Godwin, D., Martin, J.W., Marois, R., 2011. A Unified attentional
1266 bottleneck in the human brain. *Proc. Natl. Acad. Sci.* 108, 13426–13431.
1267 <https://doi.org/10.1073/pnas.1103583108>

1268 Tomov, M.S., Yagati, S., Kumar, A., Yang, W., Gershman, S.J., 2020. Discovery of hierarchical
 1269 representations for efficient planning. *PLOS Comput. Biol.* 16, e1007594.
 1270 <https://doi.org/10.1371/journal.pcbi.1007594>

1271 Tonola, C., Faroni, M., Beschi, M., Pedrocchi, N., 2023. Anytime Informed Multi-Path Replanning
 1272 Strategy for Complex Environments. *IEEE Access* 11, 4105–4116.
 1273 <https://doi.org/10.1109/ACCESS.2023.3235652>

1274 Tonola, C., Faroni, M., Pedrocchi, N., Beschi, M., 2021. Anytime informed path re-planning and
 1275 optimization for human-robot collaboration, in: 2021 30th IEEE International Conference on Robot &
 1276 Human Interactive Communication (RO-MAN). Presented at the 2021 30th IEEE International
 1277 Conference on Robot & Human Interactive Communication (RO-MAN), pp. 997–1002.
 1278 <https://doi.org/10.1109/RO-MAN50785.2021.9515422>

1279 Topalovic, U., Aghajan, Z.M., Villaroman, D., Hiller, S., Christov-Moore, L., Wishard, T.J., Stangl, M.,
 1280 Hasulak, N.R., Inman, C.S., Fields, T.A., Rao, V.R., Eliashiv, D., Fried, I., Suthana, N., 2020. Wireless
 1281 Programmable Recording and Stimulation of Deep Brain Activity in Freely Moving Humans. *Neuron*
 1282 108, 322–334.e9. <https://doi.org/10.1016/j.neuron.2020.08.021>

1283 Trier, H., O'Reilly, J.X., Spiering, L., Ma, S., Kolling, N., Rushworth, M., Scholl, J., 2023. Emotions and
 1284 individual differences shape foraging under threat. <https://doi.org/10.31234/osf.io/v6u3y>

1285 Trimmer, P.C., Houston, A.I., Marshall, J.A.R., Mendl, M.T., Paul, E.S., McNamara, J.M., 2011. Decision-
 1286 making under uncertainty: biases and Bayesians. *Anim. Cogn.* 14, 465–476.
 1287 <https://doi.org/10.1007/s10071-011-0387-4>

1288 Turesson, H., Satta, A., Domenici, P., 2009. Preparing for escape: anti-predator posture and fast-start
 1289 performance in gobies. *J. Exp. Biol.* 212, 2925–2933. <https://doi.org/10.1242/jeb.032953>

1290 von Neumann, J., Morgenstern, O., 1944. *Theory of games and economic behavior*.

1291 Walker, J.A., Ghalambor, C.K., Griset, O.L., McKENNEY, D., Reznick, D.N., 2005. Do faster starts increase
 1292 the probability of evading predators? *Funct. Ecol.* 19, 808–815. <https://doi.org/10.1111/j.1365-2435.2005.01033.x>

1294 Walters, E.T., Carew, T.J., Kandel, E.R., 1981. Associative Learning in *Aplysia*: evidence for conditioned
 1295 fear in an invertebrate. *Science*.

1296 Weintraub, I.E., Pachter, M., Garcia, E., 2020. An introduction to pursuit-evasion differential games.
 1297 Presented at the 2020 American Control Conference (ACC), IEEE, pp. 1049–1066.

1298 Weiss, K., Khoshgoftaar, T.M., Wang, D., 2016. A survey of transfer learning. *J. Big Data* 3, 9.
1299 <https://doi.org/10.1186/s40537-016-0043-6>

1300 Weller, J.E., Turner, S.P., Farish, M., Camerlink, I., Arnott, G., 2020. The Association Between Play
1301 Fighting and Information Gathering during Subsequent Contests. *Sci. Rep.* 10, 1133.
1302 <https://doi.org/10.1038/s41598-020-58063-x>

1303 Whittington, J.C.R., Muller, T.H., Mark, S., Chen, G., Barry, C., Burgess, N., Behrens, T.E.J., 2020. The
1304 Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the
1305 Hippocampal Formation. *Cell* 183, 1249–1263.e23. <https://doi.org/10.1016/j.cell.2020.10.024>

1306 Wilson, R.C., Nassar, M.R., Gold, J.I., 2010. Bayesian Online Learning of the Hazard Rate in Change-
1307 Point Problems. *Neural Comput.* 22, 2452–2476. https://doi.org/10.1162/NECO_a_00007

1308 Wiltschko, A.B., Johnson, M.J., Iurilli, G., Peterson, R.E., Katon, J.M., Pashkovski, S.L., Abaira, V.E.,
1309 Adams, R.P., Datta, S.R., 2015. Mapping Sub-Second Structure in Mouse Behavior. *Neuron*.
1310 <https://doi.org/10.1016/j.neuron.2015.11.031>

1311 Woodbury, P.B., 1986. The geometry of predator avoidance by the blue crab, *Callinectes sapidus*
1312 Rathbun. *Anim. Behav.* 34, 28–37. [https://doi.org/10.1016/0003-3472\(86\)90003-5](https://doi.org/10.1016/0003-3472(86)90003-5)

1313 Wu, C.-T., Haggerty, D., Kemere, C., Ji, D., 2017. Hippocampal awake replay in fear memory retrieval.
1314 *Nat. Neurosci.* 20, 571–580. <https://doi.org/10.1038/nn.4507>

1315 Wu, S., Éltető, N., Dasgupta, I., Schulz, E., 2023. Chunking as a rational solution to the speed–accuracy
1316 trade-off in a serial reaction time task. *Sci. Rep.* 13, 7680. [https://doi.org/10.1038/s41598-023-31500-](https://doi.org/10.1038/s41598-023-31500-3)
1317 3

1318 Wynn, M.L., Clemente, C., Nasir, A.F.A.A., Wilson, R.S., 2015. Running faster causes disaster: trade-
1319 offs between speed, manoeuvrability and motor control when running around corners in northern
1320 quolls (*Dasyurus hallucatus*). *J. Exp. Biol.* 218, 433–439. <https://doi.org/10.1242/jeb.111682>

1321 Ydenberg, R.C., Dill, L.M., 1986. The Economics of Fleeing from Predators, in: Rosenblatt, J.S., Beer, C.,
1322 Busnel, M.-C., Slater, P.J.B. (Eds.), *Advances in the Study of Behavior*. Academic Press, pp. 229–249.
1323 [https://doi.org/10.1016/S0065-3454\(08\)60192-8](https://doi.org/10.1016/S0065-3454(08)60192-8)

1324 Ye, N., Somani, A., Hsu, D., Lee, W.S., 2017. DESPOT: online POMDP planning with regularization. *J.*
1325 *Artif. Intell. Res.* 58, 231–266.

1326 Yeomans, J.S., Li, L., Scott, B.W., Frankland, P.W., 2002. Tactile, acoustic and vestibular systems sum
1327 to elicit the startle reflex. *Neurosci. Biobehav. Rev.* 26, 1–11. [https://doi.org/10.1016/s0149-](https://doi.org/10.1016/s0149-7634(01)00057-4)
1328 7634(01)00057-4

1329 Yeung, N., Summerfield, C., 2012. Metacognition in human decision-making: confidence and error
1330 monitoring. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 1310–1321. <https://doi.org/10.1098/rstb.2011.0416>
1331 You, Y., Li, W., 2016. Parallel processing of general and specific threat during early stages of
1332 perception. *Soc. Cogn. Affect. Neurosci.* 11, 395–404. <https://doi.org/10.1093/scan/nsv123>
1333 Zilberstein, S., 1996. Using Anytime Algorithms in Intelligent Systems. *AI Mag.* 17, 73–73.
1334 <https://doi.org/10.1609/aimag.v17i3.1232>
1335