

## Research report

## Active inference, stressors, and psychological trauma: A neuroethological model of (mal)adaptive explore-exploit dynamics in ecological context

Adam Linson<sup>a,b,\*</sup>, Thomas Parr<sup>c</sup>, Karl J. Friston<sup>c</sup><sup>a</sup> Faculty of Natural Sciences, University of Stirling, Stirling, UK<sup>b</sup> Faculty of Arts and Humanities, University of Stirling, Stirling, UK<sup>c</sup> Wellcome Centre for Human Neuroimaging, Institute of Neurology, University College London, London, UK

## ARTICLE INFO

## Keywords:

Stress

Fear

Aversion

Post-traumatic stress disorder (PTSD)

Psychopathology

Computational psychiatry

## ABSTRACT

This paper offers a formal account of emotional inference and stress-related behaviour, using the notion of active inference. We formulate responses to stressful scenarios in terms of Bayesian belief-updating and subsequent policy selection; namely, planning as (active) inference. Using a minimal model of how creatures or subjects account for their sensations (and subsequent action), we deconstruct the sequences of belief updating and behaviour that underwrite stress-related responses – and simulate the aberrant responses of the sort seen in post-traumatic stress disorder (PTSD). Crucially, the model used for belief-updating generates predictions in multiple (exteroceptive, proprioceptive and interoceptive) modalities, to provide an integrated account of evidence accumulation and multimodal integration that has consequences for both motor and autonomic responses. The ensuing phenomenology speaks to many constructs in the ecological and clinical literature on stress, which we unpack with reference to simulated inference processes and accompanying neuronal responses. A key insight afforded by this formal approach rests on the trade-off between the epistemic affordance of certain cues (that resolve uncertainty about states of affairs in the environment) and the consequences of epistemic foraging (that may be in conflict with the instrumental or pragmatic value of ‘fleeing’ or ‘freezing’). Starting from first principles, we show how this trade-off is nuanced by prior (subpersonal) beliefs about the outcomes of behaviour – beliefs that, when held with unduly high precision, can lead to (Bayes optimal) responses that closely resemble PTSD.

## 1. Introduction

Post-traumatic stress disorder (PTSD) is a severe mental health condition that arises following exposure to a perceived profound threat of serious injury or death. It is characterised clinically by the persistence of what are otherwise transient disturbances for most of the population, including flashbacks, hyperarousal, and avoidance behaviour [1]. Specific traumas that commonly lead to PTSD vary from transportation and industrial accidents, to street and domestic violence, to war and terror incidents. It can be a consequence of direct involvement or as a witness to these events [2].

Advances in neuroscientific and clinical research on PTSD have produced a complex picture of the condition. This complexity reflects the different perspectives on PTSD afforded by neurobiology, psychology, and behavioural science and specialisations within these areas (e.g. neuroendocrinology). The emerging field of computational psychiatry tries to integrate these perspectives in terms of computational

neuroscience [3,4]. Here, we use a computational approach based upon a generative model; namely, a lesion deficit model of how the symptoms and signs of PTSD are generated.

Specifically, this paper presents software simulations that extend previous theoretical work on a computational psychiatry model of PTSD [5]. The simulations situate this work in relation to computational ethology [6] and neuroeconomics [7]. In particular, we use a Markov decision process (MDP) model and active inference to present some scenarios – and sequential responses – that this minimal model can exhibit, on the view that, with the appropriate parameterisation, a wide variety of natural behaviours can be emulated, including approach-avoidance conflict [8]. Using MDP modelling, we aim to show that neuroethological and ecological constructs can emerge from first principles [9–11].

For background, and to convey the present aims and scope, we introduce these simulations with a brief summary of the initial theoretical model and its motivation [5]. Our overall aim in this series of papers is

\* Corresponding author at: Faculty of Natural Sciences, University of Stirling, Stirling, UK.

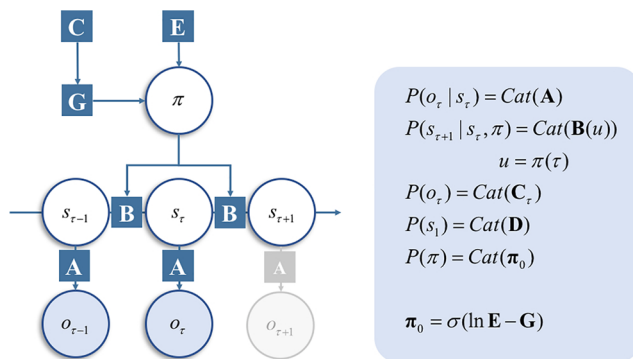
E-mail address: [adam.linson@stir.ac.uk](mailto:adam.linson@stir.ac.uk) (A. Linson).

## Box 1

### Markov decision process (MDP)

The procedure outlined here provides an intuition into the beliefs a creature (e.g. Jerry) has about how its sensory data are generated by acting on hidden states in the environment. To represent the Markov decision process used in these simulations, we refer to the factor graph below. Observed variables are shown as filled circles and unobserved variables as unfilled circles. Factors of the generative model (i.e. conditional probability distributions and prior probabilities) are shown as squares. These squares are connected to those circles containing variables that participate in the same factor. In the panel on the right, the definitions are given for each of the factors in blue squares, where *Cat* refers to the categorical distribution. The grey region of this graph indicates that the observation at the next time step is not yet available, so cannot yet be incorporated into the graph.

This structure is used to represent the *beliefs* Jerry has about how his data are caused. This means that data are never really generated from this model. Instead, this model is how Jerry draws inferences about the causes of his observed data. Despite this, it is useful to think about how we might generate data from this model to gain some intuition as to what it means. To do so, we start at the first time-step and sample a state from the categorical prior over initial states. The parameters of this prior (**D**) are simply a vector of probabilities for each alternative state. From this, we can now sample from the likelihood. This is formulated as a matrix (**A**), whose columns correspond to a state and whose rows are the alternative outcomes that may be generated. To generate an outcome, we select the column of this matrix corresponding to the state we sampled and sample an outcome from this column-vector of probabilities. It is this outcome that is available to our synthetic creature, Jerry. (For specific mappings and values, see main text and Box 2; for more general technical details, see below and [68]).



Taking a discrete time step into the future, we can sample a new state from the column of a transition matrix (**B**) associated with the state at the previous time. Crucially, the transition probabilities are conditioned upon the selected action. This means we have a separate **B**-matrix for each action. Action selection depends upon the policy, with each policy and time point associated with an action. We calculate the expected free energy (**G**) for each policy, which depends upon a vector of prior probabilities for outcomes under these policies (**C**). Combining these with a prior bias term (**E**), assumed constant across policies in this paper, we can construct a prior over policies. This uses a softmax – i.e., normalised exponential – function ( $\sigma$ ) to convert log probabilities to normalised probability distributions. Sampling from this and selecting the action that corresponds to this policy, at this time, specifies the **B**-matrix from which to sample the state for the current time step. We then sample the outcome for this time from the relevant column of the **A**-matrix. This process can be repeated for a series of discrete time steps, generating a new outcome for each time.

Importantly, the prior belief about observations only enters this graph through the expected free energy, **G**, which enters the prior over policies [68]. Policies index alternative trajectories, or sequences, of actions. In this sense, they are not time dependent, as each policy determines a sequence of actions for *all* time-points. Conversely, the actions (*u*) are time dependent. *U* is an array that specifies an action for each time step (rows) and each policy (columns). The selected action therefore depends upon the most likely policy and the action that policy implies for that time step. Action selection is technically not part of the generative model, as it relies upon the posterior distribution *Q*, obtained by inverting the model (i.e., Jerry inferring the causes of his sensations) [68]. This is an important aspect of active inference, as it underwrites the way in which the system performing inference may change the process generating its observed data.

to ground an emergent bio-behavioural phenotype in first principles, in order to produce a lesion deficit model of PTSD. By relating elements of evolutionary biology to the physics of life [9–11], we bring into view a tautology of adaptation via natural selection. Namely, any stable species – i.e., one that successfully adapts to its niche, including by shaping it – will have a biologically inherited neuronal architecture for responding to a dynamic environment that ensures species survival and reproduction.

On this basis, we define a minimal model that captures biology and behaviour conserved across species (e.g. predator avoidance). We then show how a lesion to the model produces maladaptive phenotypes. By maladaptive, we refer to what is known empirically to undermine health and reproduction, for example, a mode of persistent non-veridical perception that adversely affects heart rate. This can induce excessive metabolic costs and psychobiological pain and suffering, as described by hierarchical neurovisceral integration theories [12–14].

To ground the model in first principles, we appeal to active inference (see Box 1). This offers a framework for understanding how

organisms sample their environment to guide successful existential strategies. In other words, in the following simulations, when a world containing predators impinges on a creature's eyes and ears, those impingements lead to competing series of nerve impulses (e.g. head rotation vs. locomotion). This competition entails an implicit decision to be made: do I turn and look, run, or hold still? Or, in Bayesian terms: given the evidence at hand, what actions give me the best chance to stay out of harm's way?

Reframing well-studied phenomena in terms of active inference has beneficial metatheoretical consequences. In particular, some heterogeneous phenomena can be shown to emerge from the same underlying cause. For instance, we have theoretically motivated the idea that PTSD-relevant phenomena, including the generalisation of conditioned fear, fear extinction, and safety learning, all arise from changes in a sensory evidence accumulation threshold for a stressor inference [5]. When a creature is embedded in its ecological niche, this threshold is the crossover point from exploration to exploitation – in other words, from epistemic to pragmatic pursuits. Additionally, some superficially

similar phenomena can be shown to emerge from distinct causes. For instance, a false perception may arise from a hallucination, delusion, or a flashback. Each can be explained as a failure of sensory attenuation. However, they may each have a distinct pathogenesis, depending upon the neuromodulatory mechanisms or projections involved (e.g. within and/or between prefrontal cortex and striatum).

### 1.1. Aims and scope

Ultimately, our standpoint is that the strength of an active inference approach to PTSD lies in how it recontextualises and integrates a range of empirical evidence. It achieves this with a proposal as to why previous empirical results have obtained across a variety of methodologies and paradigms. The proposal is itself a formally grounded, computational model that rests upon inversion of a creature's internal (generative) model – generative in the sense that this specifies how consequences are generated from causes. Inversion of this model is the process of finding (inferring) the causes that best explain the observed consequences. In this paper, we describe the simplest possible version of the model that retains its core functionality in experimental simulations. This extends our pre-simulation account of the model that is anchored in extant theories of PTSD [5].

The present modelling is restricted to one aspect of our theoretical framework [5], namely, what has been widely referred to as emotional states in the psychology and neuroscience literature [15–17]. Previously, we related this state dimension of PTSD impairment to what is known as underfitting in machine learning, as follows. In brief, sparse exteroceptive sensory samples confirm a generative model of a threat, driven by strong priors, that in turn induces a neurovisceral stressor-response state.

Although beyond the present scope, we also related another aspect, emotional awareness (EA), to overfitting in machine learning. This relationship ties in with recent neuroscientific work on the interaction between emotional states and EA [18–21]. Consider that in health, EA normally follows from interoceptive inference of emotional state. For example, a visceral perception of an underlying bodily stress state could lead to the awareness that 'I feel afraid'. Under some psychopathological conditions, including anxiety and panic disorders, this causal chain may be inverted. The aberrant belief that 'I feel afraid' can result from overfitting various sensations arising from an initially non-stressed bodily state (which then becomes stressed in response).

While it is unknown how far back one must go in evolutionary history to identify the emergence of EA, it seems clear that it could not arise without a sufficiently complex neuronal architecture. For this reason, we focus the present model exclusively on emotional states. This serves to highlight the conservation of a fundamental sensorimotor survival system (the embodied neuronal architecture) that includes insects alongside mammals. In this respect, the model shares extensive commonalities with the notion of survival circuits [22], with an added emphasis on how these relate to a Bayesian account of ecological (organism-niche) dynamics.

Previous integrations of work on survival circuits with work on ethology and behavioural ecology [23,24] also share a similar focus with the present model on threat imminence (distance) and related defensive strategies [25,26]. Here too, however, our added emphasis on Bayesian ecological dynamics is relevant. Although we do not include neuroanatomical regions in the present version (this is planned for future work), we treat distance itself as an inferential construct that arises from neuronal message passing. This permits us to afford special consideration to the ambiguous relationship between auditory intensity, sound-source distance, and salience, as motivated in our initial theoretical treatment [5].

For tractability, we specify the current minimal model without the richness of human or animal awareness, healthy or impaired. Notably, a large component of PTSD research centres on more basic animal models, which similarly exclude the complexity of awareness and

related sociocultural factors. Although such complexities may indeed play a role in human PTSD, in this paper, we adhere to a minimal, biologically plausible, simulated adaptive animal model. We then deploy this as a lesion deficit model, to illustrate the emergence of a maladaptive neuro-bio-behavioral phenotype that captures some canonical animal and human PTSD symptomatology.

To present the adaptive model, we use a literal cartoon of animal brain and behaviour. This is intended as an entry point into the neuroethological implications of the active inference formal description. It is not intended to diminish the seriousness of PTSD, which negatively impacts an estimated tens of millions of people worldwide, according to recent WHO data [2]. By focussing on translational aspects of an animal (generative) model, we aim to lay the preliminary groundwork for developing concrete applications for PTSD diagnosis and treatment in humans.

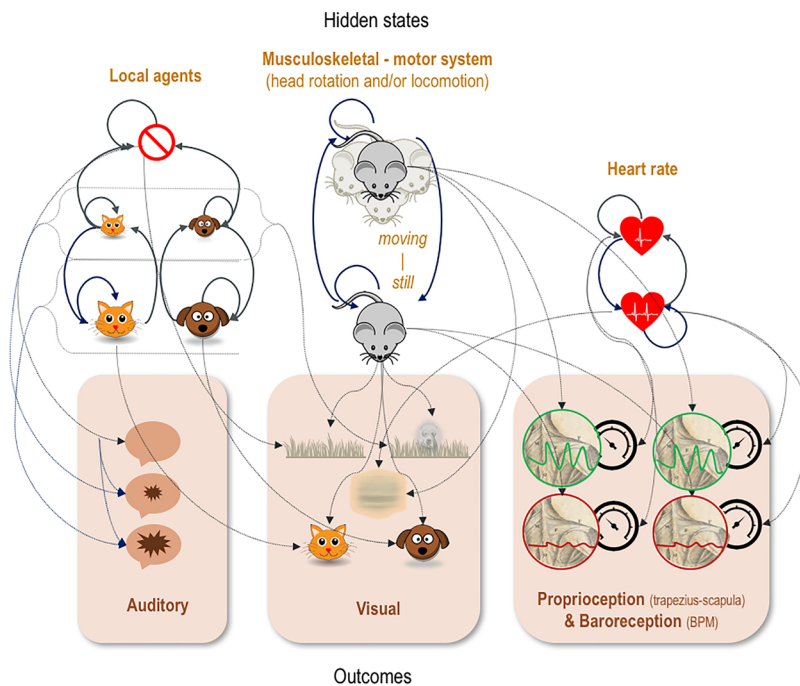
In what follows, after demonstrating the basic architecture and ensuing active inference, we then consider the emergence of adaptive and maladaptive (pathological) responses to various cues in terms of constructs related to threat and fear. We then connect these constructs to the symptomatology of PTSD, to provide a formal account of its pathophysiology. Specifically, by formulating a generative model under active inference, we can trace the implied message passing to its neurobiological substrate. Connecting the generative model to computational neuroanatomy ensures that one can predict empirical responses as measured using, for example, electrophysiological and neuroimaging data, for eventual translation into clinical and therapeutic applications [27].

This paper comprises two parts. In the first, we briefly overview active inference and the particular model used to deconstruct emotional inference in stressful situations. This model is presented in some detail, from first principles, using minimal but plausible assumptions about embodied inference. We will spend some time demonstrating the emergent phenomenology – both in terms of belief updating and accompanying the responses at both the neuronal and behavioural (motor and autonomic) levels. The second part of this paper then revisits these simulations in light of established constructs and phenomena in the literature on stress-related behaviour, with a special focus on pathology of the kind associated with PTSD.

## 2. Methods

We start by describing a generative model of active engagement with a simple world. Crucially, this model is specified in terms of biologically plausible contingencies and straightforward physical laws, without any reference to valenced states or emotional constructs. The agenda here is to demonstrate how what manifests as emotional behaviour emerges from inference based upon sensory evidence that, in itself, has no particular valence or meaning. Another important aspect of this formulation of 'emotional inference' is its multimodal nature; namely, the perceptual synthesis or inference based on exteroceptive, proprioceptive and – importantly – interoceptive sensory cues. This means the model necessarily generates sensory outcomes in multiple domains, such that model inversion constitutes multisensory integration in the service of informing emotional behaviour. In active inference, policy selection and adaptive behaviour is treated as an inference problem, much in the spirit of planning as inference [28–31].

Formally, active inference under (partially observed) Markov decision process models requires a specification of hidden states and their sensory consequences [32,33]. Once the hidden states and outcomes have been established, the parameters of the generative model determine the likelihood a particular outcome is generated by a combination of hidden states. This is usually encoded in an **A** matrix. The transitions among hidden states are parameterised in terms of **B** matrices (for each hidden state). Each hidden state is equipped with probability transition (**B**) matrices that are controlled by action, so that there is a repertoire of actions for each hidden state. Prior preferences –



**Fig. 1.** The generative model. This schematic shows the form of the generative model we assume is how Jerry makes sense of his environment. The hidden states represent the hypotheses he may appeal to in order to explain the sensory outcomes he experiences. The arrows indicate conditional probabilities, such that an arrow from one pictogram to another indicates the dependence of the second on the first. Please see the main text and [Box 2](#) for details of this model. Brief summary – environmental states ('local agents' icons): no agents, distal cat, distal dog, proximal cat, proximal dog; musculoskeletal-motor system states (see icon labels): moving, still; cardiac states (top to bottom): moderate-to-low, high; auditory outcomes (top to bottom): silence, soft sounds, loud sounds; visual outcomes (see icons): empty horizon, dot on horizon, cat shape, dog shape, blur; proprioceptive outcomes (see icons): moving or not; baroreceptive outcomes (see icons): palpable (high BPM) or not (moderate-to-low BPM).

that determine the quality or expected free energy of allowable policies – are encoded by a **C** vector over each outcome modality. Finally, beliefs about initial states are specified with **D** in a matrix for each hidden state. Equipped with this model specification (see [Box 1](#)), one can then use standard variational procedures to simulate active inference in terms of belief updating and subsequent policy selection (i.e., behaviour); for details, see [\[33\]](#).

A key aspect of this belief updating is that agents entertain (posterior) beliefs about states of the world (including the body) *and their action upon the world*; namely, the policies that determine state transitions. Recognising or inferring states of the world proceeds using standard Bayesian observer assumptions. Technically, this involves minimising a *variational free energy* (upper) bound on surprise (a.k.a. self-information). This is mathematically equivalent to maximising a (lower) bound on model evidence (a.k.a. self-evidencing) [\[34,35\]](#). This optimisation can be formulated in a biologically plausible way by associating neuronal dynamics with a gradient flow on variational free energy [\[32\]](#). The special aspect of active inference pertains to policy selection based upon inferences about which policy is being pursued (from which the next action is selected). Crucially, these inferences are based upon prior beliefs that policies minimise *expected free energy*. Minimising expected free energy can be conceptualised as selecting policies that minimise uncertainty, while leading to preferred (i.e., expected) outcomes. In other words, expected free energy combines epistemic and pragmatic imperatives in a way that dissolves the exploration-exploitation dilemma [\[36\]](#).

The particular generative model used in this paper – to demonstrate the emergence of healthy and pathological responses to safe and unsafe circumstances – aims to be as simple as possible, while being sufficiently comprehensive to account for the phenomena of interest; namely, (neuro)ethological constructs related to threat and fear, and the symptomatology and pathophysiology of PTSD [\[5\]](#). To (literally) cartoon the structure of this generative model, we will use a 'Tom and Jerry' analogy, bearing in mind that exactly the same probabilistic structure can be applied to any situation involving defensive responses to existential integrity, ranging from the pathology of predation through to psychosocial interactions.

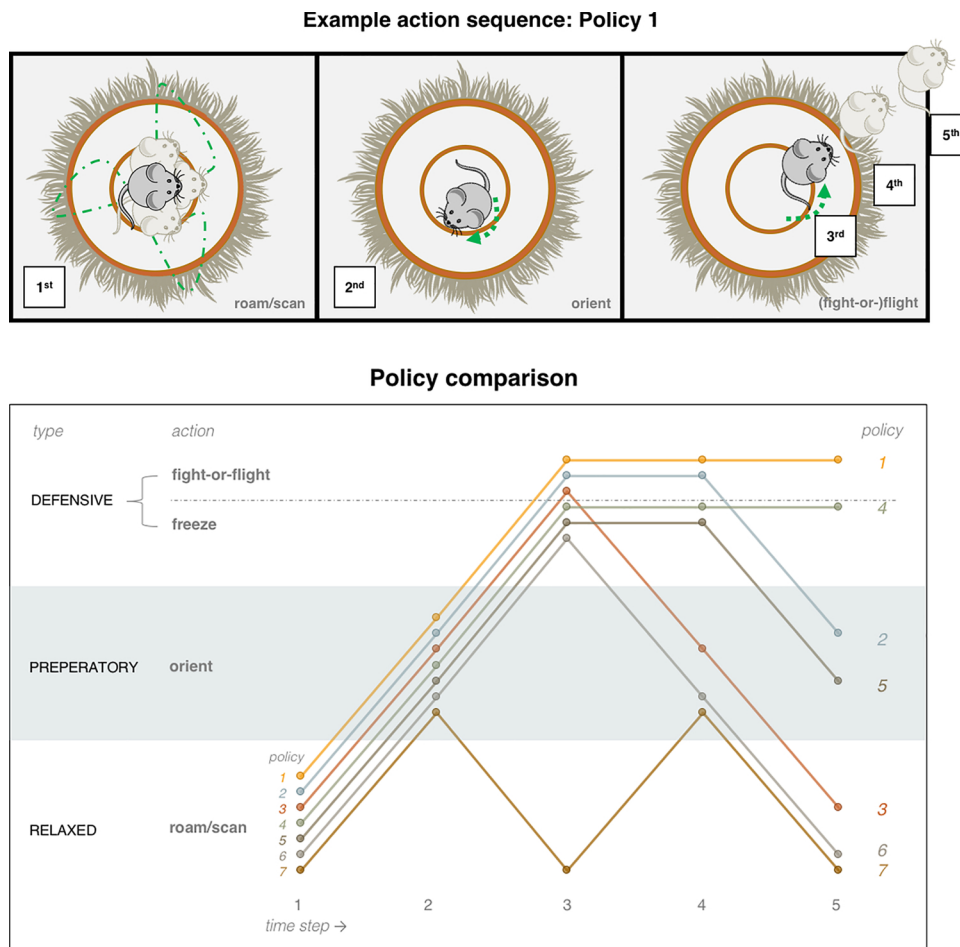
The metaphor we have in mind considers a mouse (Jerry), who must decide how to behave to avoid a cat (Tom). To set up Jerry's generative model of his environment, we start by outlining the sensory modalities

that this model must account for. Jerry's outcome modalities cover all sensory domains relevant for his inference about states of the world and policies ([Fig. 1](#)). His exteroceptive modalities are *auditory* and *visual*. The auditory modality has three levels (*silence*, *soft sounds* and *loud sounds*), and the visual modality has five levels (an *empty horizon*, a *dot on a horizon*, a *cat shape*, a *dog shape*, and a *blur*). In addition, he has a *proprioceptive* modality with a sensation that signals (self) *movement* or *not*, and a *baroreceptive* modality that reports bodily pulsations that are either *palpable* or *not*, corresponding to high or moderate-to-low beats per minute (BPM). The latter two modalities are jointly specified in four combinations. The implicit hypotheses Jerry may appeal to (sub-personally) to explain these sensory data are separated out into three sorts (factors) of hidden state. The first is the *other creatures* in his vicinity, which takes five possible values: the absence of other creatures, and either Tom or a dog, Spike, who is near or far away, respectively. The second factor is whether Jerry himself is *moving*, which he may or may not be, and the third is his own *heart rate*, which may or may not be accelerated.

With these factors and modalities in place, we can now consider the likelihood mapping (**A**) that describes how Jerry (implicitly) believes these hidden state factors give rise to their associated outcomes, and the action-dependent transitions among hidden states that underwrite beliefs about self and environmental trajectories. The likelihood of auditory outcomes depends upon the *other creatures* in Jerry's vicinity. The closer either Tom or Spike are, the louder the sounds Jerry can expect to hear. On rare occasions, a soft or loud sound may occur in the distance, even when no agents are present (e.g., a tree falling). For the visual outcome, proximal cats or dogs give rise to clear visual cat or dog forms, respectively. If they are further away, both give rise to an indistinct form on the horizon, and if they are absent, the horizon is empty. In addition, the visual outcome depends upon whether or not Jerry is moving and his heart rate. If he is moving with a high heart rate, all Jerry sees is a blur, regardless of who is in his vicinity. If he is moving without a high heart rate and there are distal creatures present, it is equiprobable that he will see a dot on the horizon or an empty horizon. Proprioceptive and baroreceptive data report, respectively, whether or not Jerry is moving, and whether or not his heart rate is rapid.

Next, we specify the dynamic structure or contingencies of the generative model: Jerry may engage in four different sorts of action, grouped into policies (sequences of actions), which have implications





**Fig. 2. Policies.** This schematic illustrates the alternative plans (or sequences of actions) that Jerry may choose to pursue. The five actions of Policy 1 are cartooned in the upper plot, which shows the following sequence: Jerry begins by (1) roaming/scanning his environment, continues on to (2) orienting, and finally, he engages in (3–5) a series of fight-or-flight actions. This sequence can be compared with the other six policies, indicated in the lower plot. To interpret this schematic, note the labels on the left that specify each action (which have been grouped into action types in the far left column). Each line is an alternative policy, with circles indicating the action at each time-step. The schematic acts as a key to interpret the labelling of the 7 alternative policies available to Jerry. This policy numbering will be adopted consistently in subsequent figures.

for all three hidden state factors. The actions that comprise these policies include: roaming/scanning; orienting; fight-or-flight; and freezing. There are seven policies (Fig. 2), three of which involve fight-or-flight actions, three of which involve freeze actions, and one which involves neither fight-or-flight nor freeze. All policies begin with roaming/scanning followed by orienting. Policies 1 and 4 continue with an unrelenting fight-or-flight or freeze, respectively. For each of these, there is a corresponding policy that turns into orienting (2 and 5), and one that reverts to roaming/scanning (3 and 6). The final policy (7) corresponds to an alternating roaming/scanning and orienting policy (see Fig. 2).

These actions contextualise the transition probabilities associated with each hidden state factor. The first transition (B) matrix allows the creatures in Jerry's vicinity to change their location according to the arrows shown in Fig. 1 (see also Box 2). For example, a nearby cat can transition probabilistically to a distant cat, but cannot change into a nearby dog. This transition depends on whether Jerry engages in his sympathetic *fight-or-flight* response, which places Jerry farther from Tom, or his parasympathetic *freeze* response, which places Tom farther from Jerry. These policy dependent contingencies manifest as Tom moving away. Jerry can similarly get away from Spike. However, when Spike wanders into the vicinity, Tom stays away. The other two policies lead to a greater probability that Tom approaches, attracted by Jerry's movement or heart rate. With respect to Jerry's intrinsic hidden states, when engaging either the *roaming/scanning* or *fight-or-flight* actions, he transitions into moving, and otherwise, if *orienting* or *freezing*, he transitions to still. When engaged in *fight-or-flight* or *orienting*, his heart rate becomes rapid, which is counteracted by the remaining two options.

Having specified the way in which data are generated from

changing states, we can now specify Jerry's prior preferences (C). He slightly disprefers seeing a blur, moderately disprefers seeing a dot on the horizon, and strongly disprefers seeing the outline of his feline predator (Tom). But, he is happy to see an empty horizon or the outline of a dog (Spike), which means Tom won't be there. In the auditory domain, Jerry prefers not to hear any sound, slightly disprefers soft sounds, and strongly disprefers loud sounds. In the jointly specified proprioceptive and baroreceptive domains, he prefers to be relaxed and moving (the result of roaming/scanning), is neutral about being excited and still (due to orienting), but disprefers the remaining combinations (the result of fight-or-flight or freezing). Note in using loaded terms like prefer and disprefer, we simply mean that the generative model includes prior probabilities about the kinds of outcomes the agent expects to encounter *a priori*.

All that remains is to set up prior beliefs about the initial states (D), which provides two alternative simulation scenarios (see Results). In the first, Jerry assumes he is in a safe location ('cat poor'), represented by flat priors. In the second, he is in an unsafe location ('cat rich'), represented by a strong prior belief that Tom is approaching from a distance. We will see below the key differences in how these scenarios affect Jerry's neurobiology and behaviour.

## 2.1. Biological grounding

Certain aspects of this generative model are motivated by known physiology. For example, policies (action sequences) are defined such that orienting always occurs before and after engaging in 'defensive' actions; namely, fight-or-flight and freezing. Orienting entails holding still, which increases visual acuity (e.g., after a head rotation to make a sound source visible). It also entails cardioacceleration (along with

**Box 2****Intact and lesioned generative model**

In relation to the intact generative model described in Fig. 1 and below, the lesion amounts to a strengthening of the prior (dis)preference for ‘I see Tom’ – implemented by changing the log probability of the value in bold from  $C_{\text{(vision)}} = [2 \ -1 \ 1 \ \mathbf{-6.25} \ -.5]$  to  $C_{\text{(vision)}} = [2 \ -1 \ 1 \ \mathbf{-10} \ -.5]$  (where the elements are, respectively, horizon, dot, dog, **cat**, blur). These values were selected to generate differential behaviour over trials. We then adjusted the values to produce plausible behavioural dynamics. Remaining preferences (for both models) include  $C_{\text{(audition)}} = [1 \ -1 \ -2]$ , i.e. no, soft, or loud sound, and  $C_{\text{(proprio-/baroreception)}} = [1 \ -1 \ -10]$ , i.e. without pressure: moving, still, and with pressure: moving, still. The safe and unsafe conditions in the experiments (see Results) entail setting the otherwise flat contextual prior (safety) to a 100 % probability that a predator (Tom) lurks in the distance ( $D_{\text{(env)}} = [0 \ 0 \ 0 \ 1 \ 0]$ ). The B matrices reflect fully controllable state transitions (see Fig. 1), with the exception of the environmental state transitions, which have the following probabilities relative to Jerry attracting or evading Tom:

$$B_{\text{(attract)}} = \begin{matrix} \begin{matrix} \text{horizon} & \text{dot} & \text{dog} & \text{cat} & \text{blur} \end{matrix} \\ \begin{matrix} \text{horizon} \\ \text{dot} \\ \text{dog} \\ \text{cat} \\ \text{blur} \end{matrix} \end{matrix} \begin{pmatrix} .25 & 0 & 0 & .1 & .3 \\ 0 & 1 & 0 & .75 & 0 \\ 0 & 0 & .5 & 0 & .4 \\ .5 & 0 & 0 & .15 & 0 \\ .25 & 0 & .5 & 0 & .3 \end{pmatrix}$$

$$B_{\text{(evade)}} = \begin{matrix} \begin{matrix} \text{horizon} & \text{dot} & \text{dog} & \text{cat} & \text{blur} \end{matrix} \\ \begin{matrix} \text{horizon} \\ \text{dot} \\ \text{dog} \\ \text{cat} \\ \text{blur} \end{matrix} \end{matrix} \begin{pmatrix} .5 & 0 & 0 & .5 & .5 \\ 0 & .1 & 0 & 0 & 0 \\ 0 & 0 & .2 & 0 & 0 \\ .25 & .9 & 0 & .5 & 0 \\ .25 & 0 & .8 & 0 & .5 \end{pmatrix}$$

Simulations were performed in Matlab using a standard routine, `spm_MDP_VB_X.m`, included with the SPM software (<http://www.fil.ion.ucl.ac.uk/spm/>).

vasodilation in skeletal muscle) to increase motor readiness (i.e. the probability of inferring a motoric policy). Physiologically, this corresponds to noradrenergic arousal and vagal and/or sympathetic activation [38–40], which has been found to be enhanced in PTSD [41].

The implications of these contingencies can be summarised as follows. When Jerry is roaming/scanning, his proprio- and baroreceptive outcomes are as he prefers (i.e., expects *a priori*). And yet, this also means he is too relaxed to adequately respond to an event of interest. Moreover, it is possible Tom is on the horizon, in or out of focal range. Thus, he must intermittently orient to prepare for motor action and subsequent epistemic foraging (which may include turning toward a distal sound source, to bring it into his visual field). Similarly, if he is moving rapidly in fight-or-flight, a blurry outcome ensues, as a proxy for degraded visual acuity – a fundamental consequence of biophysics from insects to mammals [42]. Holding still is a countermeasure, so that Jerry can assess whether or not it is safe to lower his defences (i.e. reliably resampling the scene to accumulate evidence for his posterior beliefs).

The proprioceptive outcomes were selected as a reasonably economical way to infer self-movement pertaining to forelimb use [43] and/or head rotation [44], active in roaming/scanning, and fight-or-flight, and inactive in the remaining two policies (as orienting is defined here by a contingent completed head turn, in contrast to stationary scanning). These outcomes originate from a particular head-trunk interface site found across species, homologous to the trapezius/scapula attachment in humans. This locus of proprioception is used here for neuroethological generalisation (see Discussion), given its extensive evolutionary history from 400 million year old fish up to present-day mammals [45].

In short, this minimal generative model is equipped with everything we need to account for, in terms of adaptive behaviour that has consequences in the exteroceptive, proprioceptive, and interoceptive domains. In the following, we show how the inversion of this model – using standard Bayesian message passing schemes – gives rise to sensible behaviours under the prior preferences and beliefs outlined above. We then repeat the simulations using aberrant priors to illustrate the emergence of maladaptive behaviour – in a way that reproduces the symptomatology of PTSD, and in turn points to its possible pathophysiology (see [5]). As noted above, once the generative model is

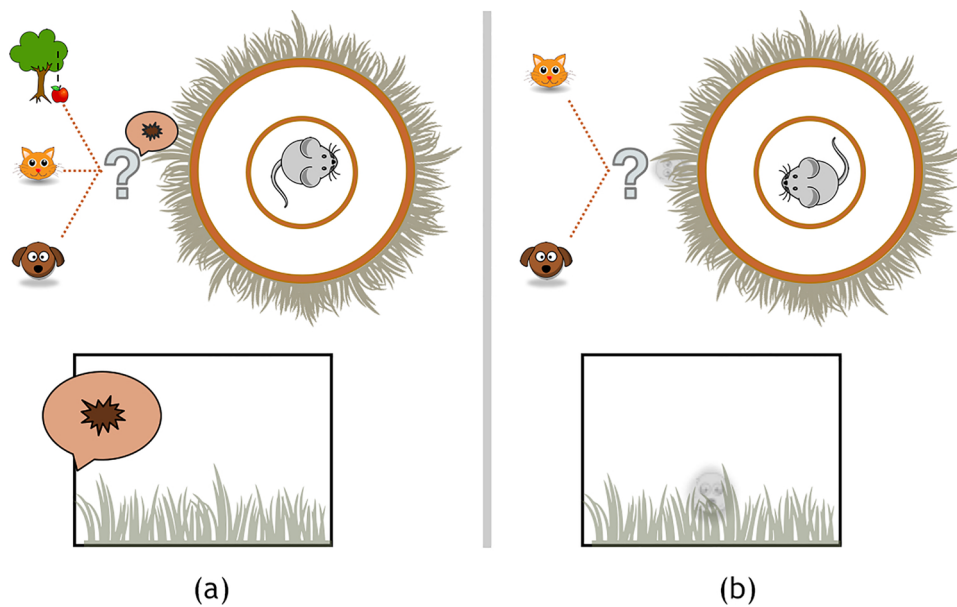
specified, the requisite (planning as) active inference can be simulated using standard variational (marginal) message passing, whose neurobiological plausibility has been established to a certain degree in several settings (see [32] for an introduction to the neuronal process theories assume in what follows).

**3. Results**

To build an intuition as to how active inference works under this sort of model, we start by describing some simple scenarios. While there are many scenarios and response sequences that this simple model can exhibit, we focus on two main narratives that best illustrate the relevant inference and behaviour. We first introduce the two scenarios, and continue by describing four variants that characterise our analysis.

In Scenario 1 (Fig. 3a), imagine Jerry is roaming/scanning, which he infers via proprioceptive movement cues and impalpable baroreceptive pulsation. For the initial exteroceptive outcomes, he sees an empty horizon, but hears a soft sound. Given his beliefs about what’s ‘out there’, he considers there to be a small chance that no agents are present, and a good chance that either Tom or Spike are approaching. The concurrent auditory and visual outcomes lead him to orient, allowing for some uncertainty resolution about the hidden state via the subsequently foraged visual outcome. If he then sees an empty horizon or dog shape, he infers he can safely resume roaming/scanning. However, if he sees a cat shape, it becomes immediately clear that a defensive policy is mandated.

Scenario 2 (Fig. 3b) also begins with Jerry roaming/scanning, and initially encountering a dot on the horizon. The dot can only be Tom or Spike, either approaching or retreating. Jerry will engage his defences if he infers Tom is approaching, in order to keep Tom at a safe distance – as he strongly disprefers seeing Tom up close. But, he also disprefers being on the defence, which he infers via proprio- and baroreceptive cues. Alternatively, he could wait and see if Spike will appear upon closer examination, but this risks that it will turn out to be Tom after all. In other words, there is an irreducible ambiguity about whether the visible dot on the horizon is Tom or Spike, so Jerry must evaluate whether to await approach, orienting for more precise state estimation – i.e. exploration (epistemic foraging) – or to instead commit to defensive action, by affording greater precision to his prior preferences –



**Fig. 3.** Scenario 1 (a), Scenario 2 (b). Lower boxes illustrate outcomes from agent POV. Upper circles and icons illustrate agent's beliefs. Specifically, the lower box in (a) portrays the visual outcome of an empty horizon, and the auditory outcome of a soft sound. The upper portion of (a) provides some intuition as to the sort of alternative hypotheses that could be used to explain this sound. It could have been generated by a cat, a dog, or could be something else (e.g. an apple dropping from a tree). The last of these hypotheses corresponds to the belief that there is a small probability of hearing a sound, even in the absence of any creatures. The lower box in (b) shows the visual outcome of a dot on the horizon, in connection with the agent evaluating counterfactual beliefs (upper portion) that the visible dot could be caused by a distal cat or distal dog. (These scenarios recur in Figs. 4 and 5).

i.e. exploitation (pragmatic foraging).

We now describe four variants for each of these scenarios. Jerry may believe that he is in a safe environment ('cat poor') or an unsafe one ('cat rich'). This is reflected, respectively, by flat priors or by a strong prior that Tom is approaching. Another variation relates to the strength of the negative prior preference (i.e., aversion) for Tom. (See Box 2 for details.) The hypothesis we test here is whether maladaptive behaviour, such as that of animal and human PTSD, can be modelled by an exaggerated prior preference. To test this, we equipped Jerry with a negative prior preference for Tom that leads to emergent adaptive behaviour; we then strengthen this negative prior preference to simulate an 'impaired' Jerry, and demonstrate anticipated emergent maladaptive behaviour. Hence, the four variants for each of the two above scenarios are: (I) unimpaired / safe; (II) unimpaired / unsafe; (III) impaired / safe; and (IV) impaired / unsafe. A subset of the simulation results for two example trials for Scenario 1, Variant I are presented in Fig. 4.

To connect this to the graphics in Fig. 4, note that the first hidden state in Fig. 4a (indicated by the cyan dot) shows no creatures in the vicinity. The outcomes generated by this state of affairs are an empty horizon and a soft sound (also indicated by cyan dots). At the second time-step, the cat is in the distance, and the same outcomes are generated. On Jerry's next action – orienting toward the sound – he sees a cat shape and hears a loud sound, as Tom has by now entered a proximal range, which Jerry (correctly) infers. This inference induces a fight-or-flight policy selection, and rapid belief updating for policies as well as states averaged under policies. As can be seen in the LFP plot, this vigorous belief-update leads to a high amplitude LFP. The selected policy results in an increase in the distance between Tom and Jerry – hence the transition from a near cat to a distant cat from the third to fourth time-steps.

The sequence in Fig. 4b shows that the dog enters distal range at the second time-step, causing a soft sound, but Jerry sees only an empty horizon. After Jerry orients toward the sound, he sees a distant dot on the horizon. However, at a distance, Jerry is unable to disambiguate between the hypotheses that the dot is caused by a cat or a dog. He has gone from wondering whether there is a cat, a dog, or nothing of interest (time-steps 1 and 2), to wondering whether there is a cat or dog (3 and 4). Only at the fifth time-step, when the dot recedes into the background, does Jerry attain precise beliefs about both the absence of a nearby creature, and about his 'better safe than sorry' freeze policy (as he has ruled out returning to a relaxed policy; see Fig. 2). This causes the high amplitude LFP plotted above (Fig. 4b). Please see the

discussion for a more in-depth analysis of this scenario.

The distributions of policy selections for each variant over 32 trials per three conditions of Scenario 1 ( $4 \times 32 \times 3$ ) and two conditions for Scenario 2 ( $4 \times 32 \times 2$ ) are depicted in Fig. 5. Note that either increasing aversion for the cat (impaired conditions), or specifying a prior belief that the cat is more probable (unsafe conditions), favour more frequent selection of more defensive policy choices, despite no change in the (hidden) states generating data. In future work modelling neuronal connectivity, it may become possible to distinguish two distinct neural correlates of these emergent PTSD phenotypes. Such a distinction could point to different clinical diagnostics and therapeutic interventions.

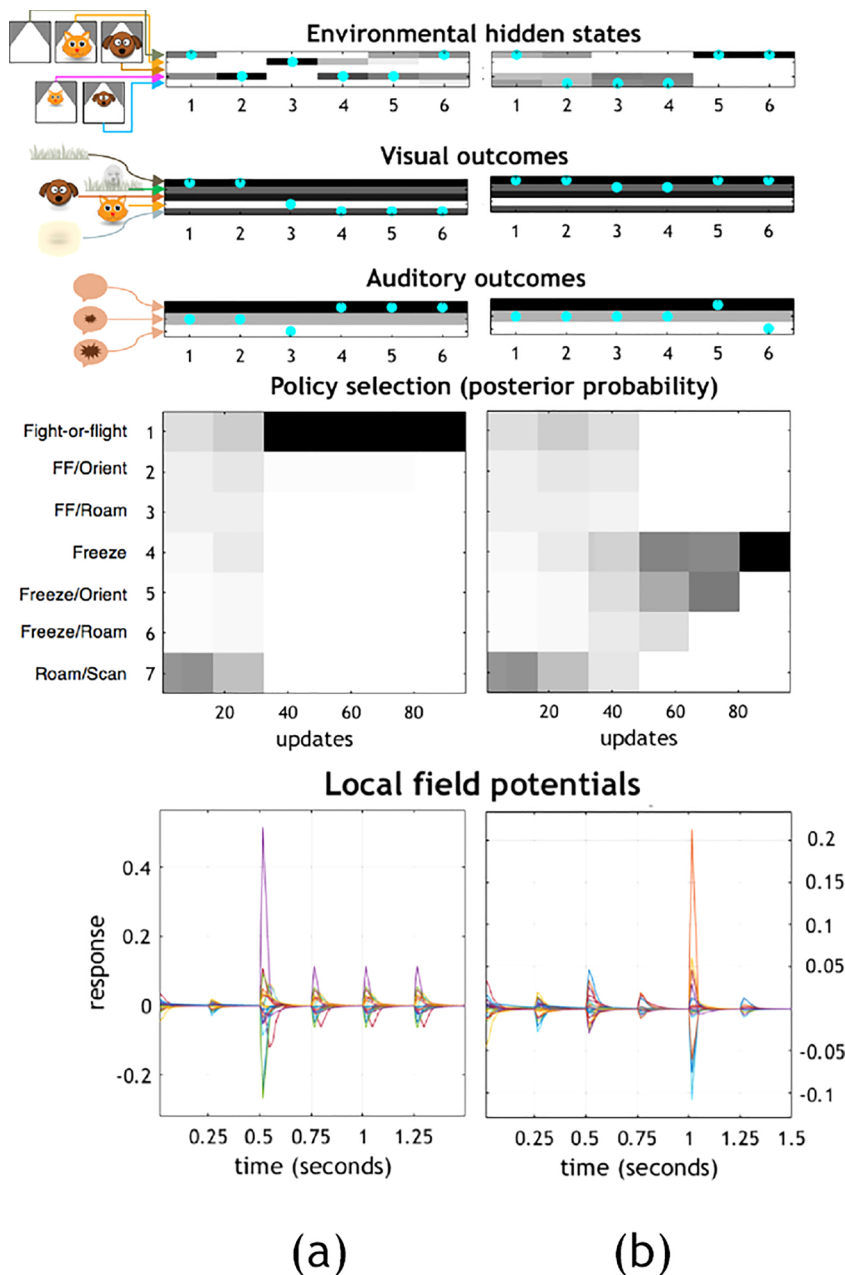
Having illustrated the characteristic behavioural responses that emerge under active inference, we now discuss how these behaviours are underwritten by emotional inference – and to what extent they can be considered a formal account of PTSD.

#### 4. Discussion

The normally intuitive relationship between 'stress' and 'stressor' differs from conventional understanding in the model on offer here. Under the generative model, causes are inferred from consequences, so an ordinarily unstressful sensory outcome (e.g. a dot on the horizon or a soft sound), under some circumstances, can lead to the same inference as an overtly stressful cue (a cat face); namely, clear evidence that a stressor is present. Given that actions affect hidden states, what is shown to be relevant above is that *stressor mitigation policies* may obtain even in the absence of stressful outcomes that would typically lead to a stressor inference. If a stressor is a destructive force (state) with a high probability of challenging existential integrity (state transition to injury or death), for reasons of natural selection, there must be a precise prior aversion to its outcomes. Crucially, on this story, *any* outcome can induce a stressor mitigation policy selection, even in the absence of an actual stressor (i.e., false inference). From this perspective, the model offers a straightforward way to link a variety of animal and human models of PTSD, including predator, trauma, stress, and psychological and biological mechanism models [46].

##### 4.1. Single trial phenomenology and neurobiology/electrophysiology

We can narrate the sequences depicted in Fig. 4 in terms of Jerry's phenomenology. Beginning with Fig. 4a, by stipulation, the scenario



**Fig. 4.** Simulated trials (columns a, b) for Scenario 1 (see Fig. 3a), Variant I (unimpaired/safe), illustrating the process of state estimation and policy selection. These plots show the trajectories of Jerry's beliefs over the course of six time steps (x-axes) with five interleaved actions; where the action sequence comprises a policy. The grayscale spectrum in bars and upper box depict probabilities, from lightest (lowest probability) to darkest (highest probability). Cyan dots indicate the true state of the world (either states or outcomes). The shading in the outcome plots indicate prior probabilities (i.e. preferences), while those in the hidden state and policy plots indicate posterior beliefs. The graphics and text to the left of the plots show alternative states of the world and policies. Interpreting the posterior beliefs as neuronal firing rates, we can use their rates of change to generate synthetic local field potentials. Note that the increase in LFP amplitude coincides with the point at which Jerry becomes confident about the policy he is pursuing.

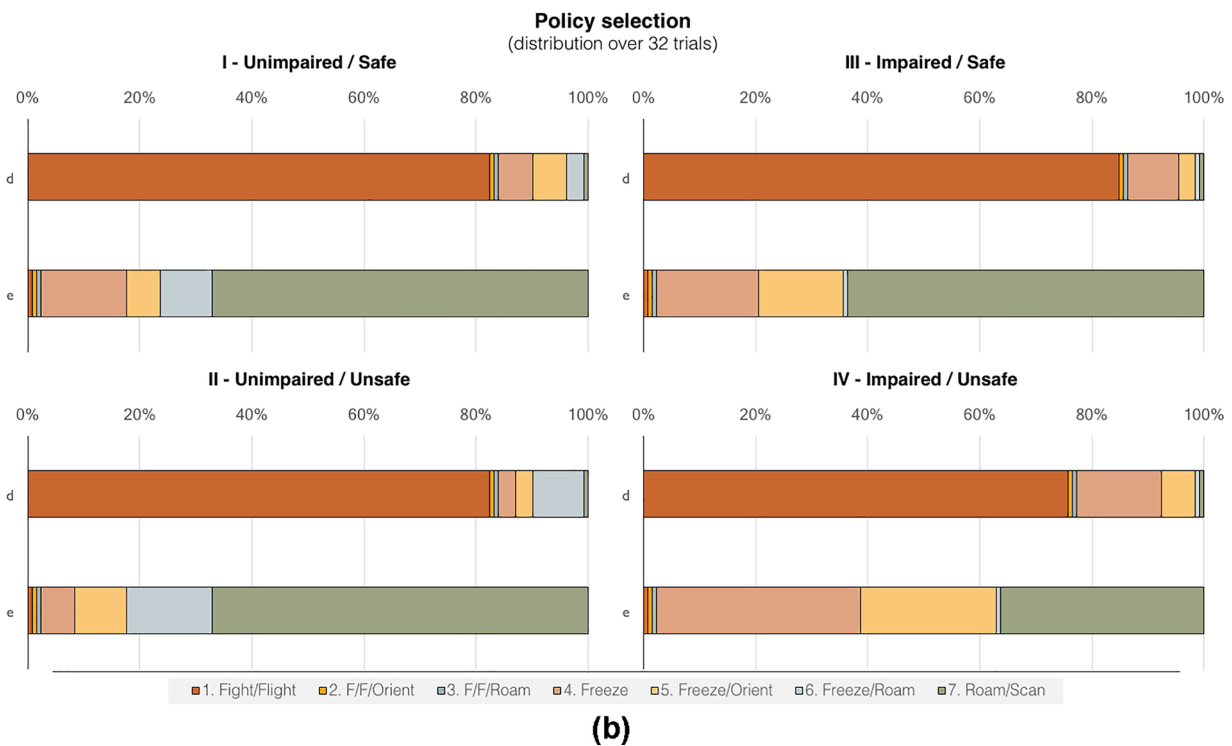
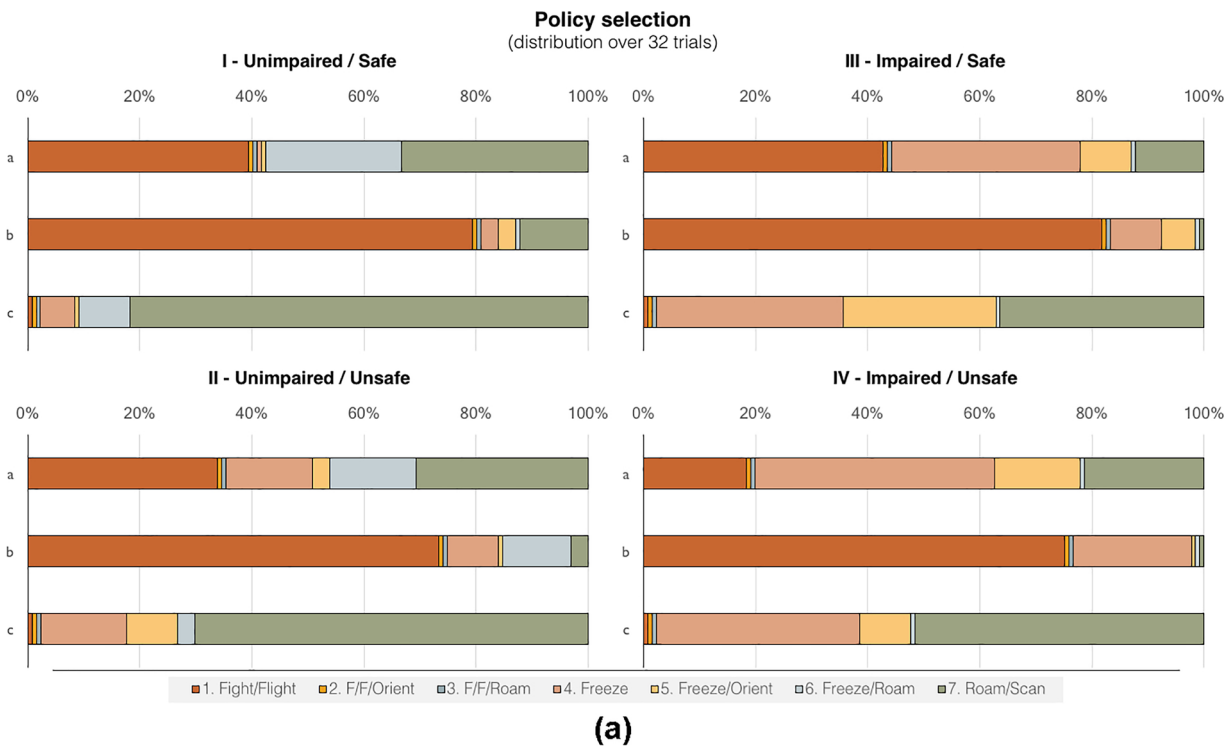
begins while Jerry is roaming/scanning, and he is compelled to continue this first action until the second time-step. In the initial state, no other creatures are in the vicinity, such that Jerry sees only an empty horizon, but hears a soft sound. In the midst of roaming/scanning, he infers that the sound may be nothing interesting (e.g. a tree falling), or may be Tom approaching, so he seriously entertains continuing to roam/scan in a carefree manner. Meanwhile, his movement is in fact attracting Tom. By the second time-step, Jerry has become certain Tom is in the distance, even though he has not yet spotted his feline foe. This leads Jerry to a second action, orienting, which prepares a defensive response by raising his heart rate (not depicted) and also locates Tom in his visual field. By the third time-step, Jerry is both certain that Tom is too close, and certain that fight-or-flight is the most apt policy (actions 3–5).

In the lower portion of Fig. 4a, we see a large LFP spike (after 500 ms.), with sufficient certainty in state estimation and defensive policy selection, and smaller spikes during subsequent state ambiguity. The relatively precise (posterior) beliefs about states and policies

correspond to movement vigour. This would plausibly correlate with the selected defensive fight-or-flight response.

The sequence in Fig. 4b has an identical set up, but reveals a different narrative and phenomenology. Under the same conditions, seeing an empty horizon and hearing a soft sound (with no creatures actually in the vicinity), Jerry considers the possibility that the sound may be nothing salient, or may be Tom or Spike distally approaching. Under the same constraints as above (roaming/scanning for action 1), at the second time-step, Jerry is still wondering whether the sound is nothing, Tom, or Spike. (In fact, Spike has wandered into the distal scene, out of view.) Jerry then orients toward the dot on the horizon (action 2), increasing his heart rate (not depicted), but this does not attract the approach of Spike, who has no interest in Jerry. In this case, Spike stays put, which keeps Tom away, but as Jerry cannot see who is in the distance, he remains concerned that it might be Tom. Thus, for action 3, Jerry holds still but has not yet fully committed to a freeze policy; he continues to entertain the possibility that it might be Spike in the distance after all, in which case he could resume carefree roaming/





**Fig. 5. (a).** This figure reports the effect of perturbing prior beliefs on the sorts of actions selected. This is shown for a range of alternative initial hidden states. This action selection distribution is constructed from 32 trials of four simulated variants (I-IV) for Scenario 1 (Initial outcomes: empty horizon visible, soft sound audible – see Fig. 3a). Initial hidden state: (a) No creatures in the vicinity; (b) Distal cat; (c) Distal dog. Each bar depicts all policies in fixed ascending order (left to right), colour-coded (see key below bars), to aid comparison across bars of the relative increase or decrease in policy selection frequency. The variants serve to reveal the key influences on policy selection. I and II (unimpaired) are defined by a moderate prior preference; III and IV (impaired) are defined by an over-strong prior preference; I and III (safe) are defined by flat priors; and II and IV (unsafe) are defined by strong priors for a distal cat. **(b).** Policy selection distribution over 32 trials of four simulated variants (I-IV) for Scenario 2 (Initial outcome: dot on horizon visible – see Fig. 3b). Initial hidden state: (d) Distal cat; (e) Distal dog. (For further details, please see Fig. 5a caption and main text.).

scanning. After holding still, whoever was in the distance (it happened to be Spike) has retreated rather than approached. Jerry is now sure no one is there, and he is correct. However, he decides roaming/scanning is not apt. He instead commits to a freeze policy (action 4–5), believing that this will ensure that Tom will not approach (better safe than sorry).

To track the above dynamics in the lower boxes of Fig. 4b, we can observe a small LFP spike (after 500 ms.), with moderate certainty about a distal agent. We also observe a large spike (after 1 s) at the elimination of remaining counterfactual policy selections that would permit Jerry's return to his preferred relaxed outcomes. This is consistent with studies of observed behaviour in non-human mammals, where a startle response proper – e.g., an abrupt, significant heart rate increase to a sudden, intense sound, with an accompanying behavioral shift to aroused and attentive, and reduced locomotor activity – is thought to be a precursor to either subsequent relaxation, fight-or-flight, or a full-fledged parasympathetic freeze response [47,48]. Based on the present simulation, this division into sub-sequences could be regarded as a separation between two freeze types. The first type is an 'a priori freeze', continuous with orienting, during which further state estimation and/or policy evaluation is performed. This is followed by the second type, a defensive response proper, in this case, an 'a posteriori freeze'. This analysis suggests an interesting consequence, namely, that a behaviourally continuous freeze state may be able to be empirically sub-divided into two phases with different underlying neurobiological and electrophysiological correlates.

#### 4.2. Toward an alternative aetiology of psychological trauma: reconceptualising fear and associative memory

With the present model, inferring a stress-related state, based on any outcome, entails inferring a policy that is expected to result in a transition to a stressor-free state. From first principles, we get three options that 'fall out' of the imperative to maintain my existential integrity by increasing the distance between me and the (inferred) danger: flight = get *farther* from danger; fight = get *further* from danger; freeze = get danger *farther* from me. Note, however, that distance is treated here as inferential construct based on relevant outcome dimensions. For instance, an increasing retinotopic surface area stimulation may reach a threshold which (from the agent's perspective) is likely to transition into the sensation of being bitten. This points to a decreasing window of opportunity for realising a successful defensive action. (This scenario illustrates inherent problems with formulations of both the 'inverse problem' and ecological 'direct perception' [9].)

Fear, although sometimes regarded as a basic emotion, can therefore instead be regarded as a *post-hoc* (observer) construct (i.e. apart from a self-awareness construct). The neurovisceral outcomes identified with fear emerge from the broader set of multimodal sensory outcomes that drive defensive action (in the service of self-maintenance), adaptively or maladaptively. Under active inference, low-level threat appraisal cannot be separated from threat response selection, as described in more detail below. This is apparent in the parallels between evolutionarily ancient neuronal survival architectures (e.g. in insects, for which a 'fear emotion' is less commonly referred to as such), and the more advanced higher-order human neuropsychology of trauma.

From this perspective, traumatic experience can be equated with a situation which blends contextually derived neurovisceral survival states (e.g. preparing for fight-or-flight) with the popular Jamesian expression of a 'blooming, buzzing confusion', which he describes as simultaneous intero- and exteroceptive assaillment. This amounts to an ecologically situated, embodied state that evinces high uncertainty in the sense of high complexity and low accuracy: many sensory outcomes are present and state estimation is confounded. During the trauma – on this story – a precise negative preference is set for an arbitrary outcome, whether or not it is in fact caused by the inflicting force, such as an intensely loud sound. This reframes 'semantic' associative memory as hypothetical inference – in other words, potentially false inference.

For instance, instead of a loud sound being merely associated with an infliction of trauma, it becomes *evidence* for this, even when trauma will not in fact be inflicted. To revisit a common example, on the present account, it can be argued that it is not the car backfiring 'cue' that reminds the combat veteran of a gunshot by traumatic memory association, as it is characterised even in the otherwise comprehensive psychobiological literature [49]. Rather, the sound of the car backfiring is never inferred as being caused by a car in the first place. Instead, it is falsely inferred as caused by a gunshot, due to a learned hidden state-sensory outcome contingency. This learned relationship could have just as easily been formed with a quiet click rather than a loud bang, or any lifespan learning that assigns (potentially aberrant) high precision to prior preference.

While we have not explicitly simulated the learning of these relationships, we implicitly assume that they have been learned in the lesioned models. A further hypothesis implied by our approach is that the learning during a traumatic experience resembles 'one shot' over-learning. At present, the simulations contrast the intact and lesioned models (for details, see Box 2). In later work, we expect an extended model will simulate how therapeutic interventions could repair the lesion, making the counterfactual space accessible again. Such repair would imply a form of learning that is closely related if not equivalent to the conventional understanding of fear extinction [50], with a different conceptual emphasis.

On an evolutionarily timescale, a 'learned' inheritance can manifest as a conserved prior aversion; e.g. for intense sound (which could be 'unlearned' during the individual lifespan). In terms of natural selection, a 'better safe than sorry' adaptation increases survival odds, and is relevant not only to subsequent reproduction but also to niche construction [51], which would conserve the adaptation (e.g., learning to avoid untraversed paths by inferring they are hostile territory, without confirmation). To take a related example, an unexpected intense sound would normally be attributed with high uncertainty to its specific source. However, an adaptation that treats loud sounds as grounds for inferring (simply) a stressor – i.e. a biologically destructive force – renders a precise inference, precipitating a vigorous avoidance response, ranging from startle to flight. Notably, this can have a maladaptive downside as well, from natural population ecology up through human social psychopathology, especially in PTSD [52].

This inferential process is evident in more basic neural architectures, such as flying insects, for which darkening fluctuations in the visual system (i.e. sensory outcomes) are sufficient to initiate escape behaviour [53]. An even more closely related example occurs in crickets that respond to (predatory) bat ultrasound, for which a mesothoracic interneuron (homologous to the human scapula-trapezius innervation site) initiates an avoidance behaviour – a burst of motor activity – when a spike-rate threshold is met; this is normally caused by a sufficiently intense acoustic perturbation [54]. As the authors point out, it is thus one in the same if the neuron is defined as a 'bat-detector' or a 'command neuron'. In our simulations, based on minimising expected free energy, both are recast as tightly coupled sensory estimation and motor policy selection to jointly secure existential continuity. Put simply, there are predicted future sensory outcomes that mandate defensive actions, whose consequences are the preferred future sensory outcomes. When uncertainty is sufficiently resolved for action, we see a (simulated) high LFP spike (Fig. 4), offering a clear empirical prediction of measurable neurobiology. That is, a high intensity sensory outcome is cast as strong evidence that facilitates low uncertainty (precise) state estimation, and consequently precise policy selection.

It is not intrinsically important whether the stimulus intensity is high or low, but rather, what matters is that it correlates with a precise state estimation that precipitates avoidance. This puts a different spin on why olfaction is such an effective modality with respect to detecting predators. Their scent (outcome) is unambiguous, so by strongly dispreffering it, there will always be a relatively sure bet that the predator (state) can be avoided. That is, natural selection conserves the precise

negative preference for particular olfactory outcomes, although other trade-offs come into play with purely chemosensory assessment [55]. This can be contrasted with vision, which has greater ambiguity – you might mistakenly stay put and become lunch for a camouflaged or distant, fast-approaching predator – and auditory intensity, which has even greater ambiguity, given harmless loud events and dangerous quiet ones in the same frequency range.

In this light, from an information-theoretic perspective, in contrast to conventional animal communication and signalling theories, active inference allows for the emergence of a neuroethological and ecological story from first principles: it is a property of the generative model that multimodal integration lowers uncertainty or entropy. Empirical studies confirm that combining auditory and visual cues increases the robustness, precision, and discrimination of perceptual evidence relevant to decision-making [56–58]. This is accounted for in the present simulations when combined auditory and visual cues resolve uncertainty about the presence of agents in the vicinity. Hence, natural selection favours multimodal integration, and conserves a ‘better safe than sorry’ fallback under greater uncertainty, such as when only audition or vision is available.

In short, depending on the neuronal architecture and character of the initial sensory evidence, agents bypass the question ‘what is it?’ and get straight to the point: ‘is there sufficient evidence that my integrity is in jeopardy?’. When the answer is ‘yes’, this implies evidence in favour of a pragmatic course of action – realising the preferred outcome by decreasing the sound intensity via fleeing – which will be a vigorous action in virtue of its relative certainty [59]; i.e., a policy selection that minimises expected free energy by assigning high precision to prior preferences (Fig. 4a). In this light, the adaptive/maladaptive and healthy/pathological distinctions relate simply to the sufficiency threshold of evidence that my integrity is in jeopardy. For some individuals, particular traumatic experiences would appear to semi-permanently lower this threshold such that it is easily met with imprecise evidence, as illustrated by our lesion deficit model (see Results and Box 2).

If this threshold for possible danger is high enough, the intense auditory outcome might be initially met with a vigorous head turn, to perform quick epistemic foraging for uncertainty-resolving outcomes, for example, by foveating the source location [60]. The decision to select epistemic over pragmatic foraging – amounting to exploration outweighing exploitation – is based on an evaluation of counterfactual integrity-preserving actions. Namely, turning toward the source rather than fleeing immediately would conserve metabolic resources that would be wasted on an unnecessary diversion to skeletal muscles.<sup>1</sup> Indeed, such unnecessary diversion takes place in maladaptive (unmerited) hyperarousal, discussed below.

#### 4.3. Distributions of policy selections over 4 variants, 5 conditions, 32 trials per condition: emergent hyperarousal and hypervigilance

Both the epistemic (orienting) and pragmatic (defensive) responses described above are consistent with the startle reflex, which can be contextualised within an ancient evolutionary sensorimotor architecture comprised of orienting and defensive vagal and/or sympathetic response mediation of the autonomic nervous system, especially the heart [61]. The exaggerated startle reflex and corresponding hyperarousal found in PTSD patients [62] is modelled here as unduly high precision on a prior preference that therefore more frequently evokes the pragmatic response. This emergent hyperarousal is reflected in Fig. 5 as a decrease in light blue bands in all scenarios and conditions,

<sup>1</sup> Of course, if the embodied brain is there as a product of conserving metabolic reserves and the attending evolutionary imperatives, it need not know this; it must merely inherit precise (subpersonal) priors that amount to ‘I am not the sort of creature that moves unnecessarily’.

and in Scenario 1 (Fig. 5a) under presumed safe conditions (I and III) as a substantial decrease in green bands – and a corresponding substantial increase in light orange and light yellow bands. The impaired responses under presumed safety especially closely mimics PTSD-related hyperarousal, as the defensive policies obtain in situations for which the unimpaired agents frequently find no grounds for raising (or maintaining raised) defences.

A further nuance captured by this model relates to startled freezing behaviour. Consider that for an epistemically beneficial head turn (i.e., orienting), a possible consequence is the encounter of an aversive outcome and corresponding proximal threat inference, that in turn precipitates a pragmatic response [63]. In this case, vagal and/or parasympathetic lowering of the heart rate has a high probability of being invoked (‘freezing’). This action realises the pragmatic aim of maintaining a corresponding low probability of being detected by the threat [64], i.e. reducing the chance of a strongly dispreferred proximal threat outcome. However, freezing also conflicts with the aim (preference) of avoiding vulnerable self-defensive state outcomes. When the latter aim wins out, heuristically, the stage is set for further epistemic foraging [65], which approximates the constructs of being courageous and curious over fearful and withdrawn.

The ability to transition out of a vulnerable self-defensive state therefore requires having low enough precision on the prior aversion, which can be interpreted as cognitive flexibility, associated with post-traumatic recovery [[66] (see also [67]). An impaired agent – characterised here with excessively high precision on the prior aversion – may therefore become stuck in what closely mimics a hypervigilance behaviour (Fig. 5b), a common PTSD symptom. Note that in presumed safety (I and III), and in the absence of actual danger (e), the contrast between the unimpaired (I) and impaired agent (III) is such that the former is able to resume roaming/scanning after a momentary freeze (light blue band) more frequently than the latter, who instead goes from freezing to aroused orienting (light yellow band). In presumed unsafe conditions (II and IV), in both actual danger (d) and its absence (e), the impaired agent (IV) also shows a substantial increase (approximately 4-fold) in ‘hard’ freezing (light orange band) compared to the unimpaired agent (II).

## 5. Summary and conclusions

We have presented a generative model of adaptive/healthy and maladaptive/pathological behaviour, grounded in an underlying evolutionary psychobiology. This computational approach offers a novel characterisation of PTSD based on first principles, rather than on conventional constructs such as ‘fear’ and ‘associative memory’, while remaining consistent with empirical findings linked to these constructs. The primary feature of the model is its use of overly precise prior preferences/aversions to model aberrant psychophysiological responses, where these precise priors are conserved by natural selection.

By treating the stressor (hidden state) inference as more fundamental than outcomes providing evidence for stress, we have proposed the following schema, that unifies standard active inference, neuroethological constructs, and a PTSD phenotype. Namely, in specific contexts, exploration (epistemic foraging) is invoked for resolving uncertainty about the presence of a stressor. A high negative preference (set by evolution or learning) for a stressor-dependent outcome rapidly leads to exploitation (pragmatic foraging) before further state estimation. A ‘fearful’ or ‘aversive’ stimulus is thereby recast as an aversive outcome related to a stressor inference, with high precision assigned to the aversion.

Two common PTSD associated phenomena, the ‘generalisation of conditioned fear’ (a.k.a., an inability for disambiguation or discrimination of a fearful stimulus) and ‘fear extinction’ (a.k.a., safety learning) are also recast as two sides of the same coin – a dynamic balancing act between exploration and exploitation. Specifically, if excessively high precision is assigned to prior preference, this will limit

the exploratory resolution of uncertainty regarding the inferred hidden state for a potentially stressor-dependent cue. Conversely, when this precision on preference is lowered sufficiently, the exploitative drive gives way to exploratory imperatives. In other words, when minimal evidence is deemed sufficient for inferring a stressor is present, a reaction resembling fear is induced (exploitation). When minimal evidence is appropriately regarded as insufficient for a stressor inference, depending on the agent's focus, a reaction resembling curiosity is induced (exploration), or one of indifference to a stressor hypothesis.

On this understanding, policies leading to a stressor-absent state are invoked for reducing uncertainty relative to precise prior preferences. These stressor-mitigation policies induce continuous arousal, and may invoke policies associated with movement vigour (e.g. fight-or-flight). Thus, startle is modelled as an aroused, vigorous response to a sensory outcome (e.g. sound), which may or may not transition into a defensive response, depending on a number of conditions. The conditions pertain to the fact that further exploration increases risk – that is, it raises the expected free energy of allowable policies because it lowers the probability of escaping the state that generates aversive outcomes. This points to a counterfactual policy evaluation delay, characterised here as an ‘*a priori*’ freeze, which differs from a defensive freeze, despite their behavioural identity. Moreover, this characterisation of startle permits us to model exaggerated startle, a common PTSD symptom, as arising from excessively precise prior preferences.

Finally, the model provides an explanation for why we should expect to find hyperarousal and hypervigilance in PTSD patients, as is the case. The former relates here to excessive (salience-related) motor policy readiness – that is, when motoric policies are the most probable – due to such policies having the lowest expected free energy (lowest risk) for an active transition to a stressor-free state. This can emerge independently, or in conjunction with hypervigilance. The latter is framed as a sub-variety of excessive epistemic foraging, in which the predominant operative hypothesis (for which further evidence is sought) relates to the pragmatic avoidance of highly dispreferred outcomes.

In conclusion, active inference accounts for a wide range of evidence that, in some cases, has been disclosed within narrow disciplinary silos. The confinement of research to inherited constructs can at times suggest an impasse has been reached. However intuitive these predominant constructs may have become, rather than accept them *a priori*, we can reproduce construct-related phenomena as emergent from first principles. If, on this basis, synthetic PTSD can be shown to have adequate neuro-bio-behavioural grounding – a project that will continue with future work – it should be possible to use this computational approach to build a bridge between basic research in psychobiology and translational clinical diagnostic and therapeutic insights to improve health and wellbeing.

## Author's contribution

All authors contributed to all aspects of the published work.

## Funding

TP is supported by the Rosetrees Trust (Award Number 173346). KJF is a Wellcome Principal Research Fellow (Ref: 088130/Z/09/Z).

## References

- [1] J.E. Sherin, C.B. Nemeroff, Post-traumatic stress disorder: the neurobiological impact of psychological trauma, *Dialogues Clin. Neurosci.* 13 (3) (2011) 263.
- [2] R.C. Kessler, S. Aguilar-Gaxiola, J. Alonso, C. Benjet, E.J. Bromet, G. Cardoso, L. Degenhardt, G. de Girolamo, R.V. Dinolova, F. Ferry, S. Florescu, O. Gureje, J.M. Haro, Y. Huang, E.G. Karam, N. Kawakami, S. Lee, J.-P. Lepine, D. Levinson, F. Navarro-Mateu, B.-E. Pennell, M. Piazza, J. Posada-Villa, K.M. Scott, D.J. Stein, M. Ten Have, Y. Torres, M.C. Viana, M.V. Petukhova, N.A. Sampson, A.M. Zaslavsky, K.C. Koenen, Trauma and PTSD in the WHO world mental health surveys, *Eur. J. Psychotraumatol.* 8 (Suppl 5) (2017) 1353383.
- [3] P.R. Montague, R.J. Dolan, K.J. Friston, P. Dayan, Computational psychiatry, *Trends Cogn. Sci.* 16 (1) (2012) 72–80.
- [4] T. Mäki-Marttunen, T. Kaufmann, T. Elväsahagen, A. Devor, S. Djurovic, L.T. Westlye, M.-L. Linne, M. Rietschel, D. Schubert, S. Borgwardt, M. Efrim-Budisteanu, F. Bettella, G. Haines, E. Hagen, S. Naess, T.V. Ness, T. Moberget, C. Metzner, A.G. Edwards, M. Fyhn, A.M. Dale, G.T. Einevoll, O.A. Andreassen, Biophysical psychiatry—how computational neuroscience can help to understand the complex mechanisms of mental disorders, *Front. Psychiatry* 10 (2019) 534.
- [5] A. Linson, K. Friston, Reframing PTSD for computational psychiatry with the active inference framework, *Cogn. Neuropsychiatry* 24 (5) (2019) 347–368.
- [6] D.J. Anderson, P. Perona, Toward a science of computational ethology, *Neuron* 84 (1) (2014) 18–31.
- [7] K.K. Watson, M.L. Platt, Neuroethology of reward and decision making, *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 363 (1511) (2008) 3825–3835.
- [8] H.A.J. Metz, H. Dieneske, G. de Jonge, F.A. Putters, Continuous-time markov chains as models for animal behaviour, *Bull. Math. Biol.* 45 (4) (1983) 643–658.
- [9] A. Linson, A. Clark, S. Ramamoorthy, K. Friston, The active inference approach to ecological perception: general information dynamics for natural and artificial embodied cognition, *Front. Robot. AI* 5 (2018).
- [10] M.J.D. Ramstead, P.B. Badcock, K.J. Friston, Answering Schrödinger's question: a free-energy formulation, *Phys. Life Rev.* 24 (2018) 1–16.
- [11] J. Daunizeau, A plea for “variational neuroethology”: comment on “Answering Schrödinger's question: a free-energy formulation” by M.J. Desormeau Ramstead et al., *Phys. Life Rev.* 24 (2018) 56–58.
- [12] R. Smith, J.F. Thayer, S.S. Khalsa, R.D. Lane, The hierarchical basis of neurovisceral integration, *Neurosci. Biobehav. Rev.* 75 (2017) 274–296.
- [13] A.K. Seth, K. Friston, Active interoceptive inference and the emotional brain, *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 371 (1708) (2016).
- [14] A.P. Owens, M. Allen, S. Ondobaka, K.J. Friston, Interoceptive inference: from computational neuroscience to clinic, *Neurosci. Biobehav. Rev.* 90 (2018) 174–183.
- [15] K.R. Scherer, The dynamic architecture of emotion: evidence for the component process model, *Cogn. Emot.* 23 (7) (2009) 1307–1351.
- [16] R. Smith, R.D. Lane, The neural basis of one's own conscious and unconscious emotional states, *Neurosci. Biobehav. Rev.* 57 (2015) 1–29.
- [17] R. Smith, R.D. Lane, Unconscious emotion: a cognitive neuroscientific perspective, *Neurosci. Biobehav. Rev.* 69 (2016) 216–238.
- [18] J. Panksepp, R.D. Lane, M. Solms, R. Smith, Reconciling cognitive and affective neuroscience perspectives on the brain basis of emotional experience, *Neurosci. Biobehav. Rev.* 76 (2017) 187–215.
- [19] R. Smith, W.D.S. Killgore, R.D. Lane, The structure of emotional experience and its relation to trait emotional awareness: a theoretical review, *Emotion* 18 (5) (2018) 670–692.
- [20] L.F. Barrett, How Emotions Are Made: The Secret Life of the Brain, Pan Macmillan, 2017.
- [21] R. Smith, R.D. Lane, T. Parr, K.J. Friston, Neurocomputational mechanisms underlying emotional awareness: insights afforded by deep active inference and their potential clinical relevance, *Neurosci. Biobehav. Rev.* (2019).
- [22] J. LeDoux, Rethinking the emotional brain, *Neuron* 73 (4) (2012) 653–676.
- [23] D. Mobbs, J.J. Kim, Neuroethological studies of fear, anxiety, and risky decision-making in rodents and humans, *Curr. Opin. Behav. Sci.* 5 (2015) 8–15.
- [24] D. Mobbs, P. Petrovic, J.L. Marchant, D. Hassabis, N. Weiskopf, B. Seymour, R.J. Dolan, C.D. Frith, When fear is near: threat imminence elicits prefrontal-periaqueductal gray shifts in humans, *Science* 317 (5841) (2007) 1079.
- [25] R.J. Blanchard, D.C. Blanchard, Anti-predator Defense as Models of Animal Fear and Anxiety, Fear and Defence, Harwood Academic Publishers, Amsterdam, Netherlands, 1990, pp. 89–108.
- [26] N. McNaughton, P.J. Corr, A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance, *Neurosci. Biobehav. Rev.* 28 (3) (2004) 285–305.
- [27] T. Parr, K.J. Friston, The anatomy of inference: generative models and brain structure, *Front. Comput. Neurosci.* 12 (2018) 90.
- [28] R. Kaplan, K.J. Friston, Planning and navigation as active inference, *Biol. Cybern.* (2018).
- [29] C.L. Baker, R. Saxe, J.B. Tenenbaum, Action understanding as inverse planning, *Cognition* 113 (3) (2009) 329–349.
- [30] H. Attias, Planning by probabilistic inference, *Proc. of the 9th Int. Workshop on Artificial Intelligence and Statistics* (2003).
- [31] M. Botvinick, M. Toussaint, Planning as inference, *Trends Cogn. Sci.* 16 (10) (2012) 485–488.
- [32] K. Friston, T. FitzGerald, F. Rigoli, P. Schwartenbeck, G. Pezzulo, Active inference: a process theory, *Neural Comput.* 29 (1) (2017) 1–49.
- [33] K.J. Friston, T. Parr, B. de Vries, The graphical brain: belief propagation and active inference, *Netw. Neurosci.* 0 (0) (2017) 1–34.
- [34] J. Hohwy, The self-evidencing brain, *Noûs* 50 (2) (2016) 259–285.
- [35] J. Winn, C.M. Bishop, Variational message passing, *J. Mach. Learn. Res.* 6 (2005) 661–694.
- [36] K. Friston, F. Rigoli, D. Ognibene, C. Mathys, T. FitzGerald, G. Pezzulo, Active inference and epistemic value, *Cogn. Neurosci.* 6 (4) (2015) 187–214.
- [37] R.A.L. Dampney, Chapter 28 - central mechanisms generating cardiovascular and respiratory responses to emotional stress, in: G. Fink (Ed.), *Stress: Physiology, Biochemistry, and Pathology*, Academic Press, 2019, pp. 391–402.
- [38] S.M. Hilton, The defence-arousal system and its relevance for circulatory and respiratory control, *J. Exp. Biol.* 100 (1) (1982) 159.
- [39] C.W. Berridge, Noradrenergic modulation of arousal, *Brain Res. Rev.* 58 (1) (2008) 1–17.
- [40] G. Ronzoni, A. del Arco, F. Mora, G. Segovia, Enhanced noradrenergic activity in



- the amygdala contributes to hyperarousal in an animal model of PTSD, *Psychoneuroendocrinology* 70 (2016) 1–9.
- [42] M.F. Land, Motion and vision: why animals move their eyes, *J. Comp. Physiol. A* 185 (4) (1999) 341–352.
- [43] D. Roberts, Structure and Function of the Primate Scapula, Primate Locomotion, Academic Press, New York, 1974, pp. 171–200.
- [44] J.C. DeToledo, N.J. David, Innervation of the sternocleidomastoid and trapezius muscles by the accessory nucleus, *J. Neuroophthalmol.* 21 (3) (2001) 214–216.
- [45] T. Matsuoaka, P.E. Ahlberg, N. Kessaris, P. Iannarelli, U. Dennehy, W.D. Richardson, A.P. McMahon, G. Koentges, Neural crest origins of the neck and shoulder, *Nature* 436 (7049) (2005) 347–355.
- [46] M.A. Matar, J. Zohar, H. Cohen, Translationally relevant modeling of PTSD in rodents, *Cell Tissue Res.* 354 (1) (2013) 127–139.
- [47] J.C. Talling, N.K. Waran, C.M. Wathes, J.A. Lines, Behavioural and physiological responses of pigs to sound, *Appl. Anim. Behav. Sci.* 48 (3) (1996) 187–201.
- [48] J.L. Lanier, T. Grandin, R.D. Green, D. Avery, K. McGee, The relationship between reaction to sudden, intermittent movements and sounds and temperament, *J. Anim. Sci.* 78 (6) (2000) 1467–1474.
- [49] R.K. Pitman, A.M. Rasmusson, K.C. Koenen, L.M. Shin, S.P. Orr, M.W. Gilbertson, M.R. Milad, I. Liberzon, Biological studies of post-traumatic stress disorder, *Nat. Rev. Neurosci.* 13 (2012) 769.
- [50] M.E. Bouton, Context and behavioral processes in extinction, *Learn. Mem.* 11 (5) (2004) 485–494.
- [51] A. Constant, M. Ramstead, J.D.S. Veissière, P.L.J. Campbell, O.K. Friston, J.A. variational approach to niche construction, *J. R. Soc. Interface* 15 (141) (2018) 20170685.
- [52] L.Y. Zanette, M. Clinchy, Predator–prey Interactions: Integrating Fear Effects, *APA Handbook of Comparative Psychology: Basic Concepts, Methods, Neural Substrate, and Behavior Vol. 1* American Psychological Association, Washington, DC, US, 2017, pp. 815–831.
- [53] M.H. Holmqvist, M.V. Srinivasan, A visually evoked escape response of the housefly, *J. Comp. Physiol. A* 169 (4) (1991) 451–459.
- [54] R. Hoy, T. Nolen, P. Brodfuehrer, The neuroethology of acoustic startle and escape in flying insects, *J. Exp. Biol.* 146 (1) (1989) 287.
- [55] L.B. Kats, L.M. Dill, The scent of death: chemosensory assessment of predation risk by prey animals, *Ecoscience* 5 (3) (1998) 361–394.
- [56] V. Best, E.J. Ozmeral, B.G. Shinn-Cunningham, Visually-guided attention enhances target identification in a complex auditory scene, *JARO* 8 (2) (2007) 294–304.
- [57] R. Lewis, U. Noppeney, Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas, *J. Neurosci.* 30 (37) (2010) 12329.
- [58] U. Noppeney, D. Ostwald, S. Werner, Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex, *J. Neurosci.* 30 (21) (2010) 7434.
- [59] I. Opris, M. Lebedev, R. Nelson, Motor planning under unpredictable reward: modulations of movement vigor and primate striatum activity, *Front. Neurosci.* 5 (2011) 61.
- [60] R.S. Heffner, Primate hearing from a mammalian perspective, *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* 281A (1) (2004) 1111–1122.
- [61] B.A. Campbell, G. Wood, T. McBride, Origins of orienting and defensive responses: an evolutionary perspective, in: P.J. Lang, R.F. Simons, M. Balaban (Eds.), *Attention and Orienting: Sensory and Motivational Processes*, Lawrence Erlbaum, Mahwah, NJ, 1997, pp. 41–67.
- [62] N. Pole, The psychophysiology of posttraumatic stress disorder: a meta-analysis, *Psychol. Bull.* 133 (5) (2007) 725–746.
- [63] K. Friston, F. Rigoli, D. Ognibene, C. Mathys, T. Fitzgerald, G. Pezzulo, Active inference and epistemic value, *Cogn. Neurosci.* 6 (4) (2015) 187–214.
- [64] K. Roelofs, Freeze for action: neurobiological mechanisms in animal and human freezing, *Philos. Trans. R. Soc. B: Biol. Sci.* 372 (1718) (2017).
- [65] M.B. Mirza, R.A. Adams, C. Mathys, K.J. Friston, Human visual exploration reduces uncertainty about the sensed world, *PLoS One* 13 (1) (2018) e0190429.
- [66] R. Yehuda, J. LeDoux, Response variation following trauma: a translational neuroscience approach to understanding PTSD, *Neuron* 56 (1) (2007) 19–32.
- [67] L. Sinapayen, A. Masumori, T. Ikegami, Learning by stimulation avoidance: a principle to control spiking neural networks dynamics, *PLoS One* 12 (2) (2017) e0170388.
- [68] T. Parr, K.J. Friston, Generalised free energy and active inference, *Biol. Cybern.* (2019).

## Glossary

*In Bayesian statistics and machine learning, several common terms have technical meanings. This glossary defines the way in which we use key terms in the current article*

**Free energy:** an information theoretic measure that bounds (is greater than) the surprise on sampling some data, given a generative model.

**Expected free energy:** under active inference, expected free energy is the variational free energy expected under (posterior predictive) beliefs about outcomes. This scores the evidence for plausible policies based on outcomes that have yet to be observed.

**Entropy:** the average surprise of outcomes sampled from a probability distribution or density. A density with low entropy means, on average, the outcome is relatively predictable. High entropy denotes unpredictability and uncertainty.

**Surprise, surprisal, or self-information:** the negative log-probability of an outcome. An improbable outcome is, therefore, surprising. Negative surprise is the same as log evidence; namely, the logarithm of Bayesian model evidence.

**Bayesian surprise:** a measure of salience based on the divergence between the posterior and prior probability densities. It measures the information gain obtained by updating the priors to posteriors.

**Generative model:** a probabilistic model that generates consequences (i.e., data) from their causes (i.e., model parameters). A generative model is also known as a forward model and is usually specified in terms of the likelihood of getting some data given their causes (parameters of a model) and priors on the parameters.

**Prior:** the probability distribution or density over the causes of data that encode beliefs about those causes prior to observing the data.

**Empirical prior:** priors that are induced by hierarchical models; they provide constraints on the recognition density in the usual way but depend on the data.

**Posterior density:** the probability distribution over causes or model parameters, given some data; i.e., a probabilistic mapping from observed consequences to causes. In Bayesian inference, the prior is updated—on the basis of observations—to become a posterior, according to Bayes rule.

**Model evidence:** in Bayesian statistics, the model evidence is the probability that observed data were generated by a particular generative model. The negative logarithm of model evidence is surprise or self-information in information theory.

**Policy:** a trajectory or sequence of actions that is determined for all time-points in a series. Under the Markov decision process (MDP) scheme used here, a selected action depends upon the most likely policy.

**Hidden state:** a latent state of the self or world that is inferred; (subpersonal) beliefs about hidden states can be revised to match observed evidence, or new evidence can be gathered to make observations consistent with beliefs about hidden states.

**Outcomes:** under active inference, outcomes refer to sensory evidence, typically defined for a given sensory modality.