The Depressed Brain: An Evolutionary Systems Theory

Paul B. Badcock[1,2,3]

E-mail: pbadcock@unimelb.edu.au; Ph: +61450211976

Christopher G. Davey[1,3]

E-mail: c.davey@unimelb.edu.au; Ph: +61 403 058 343

Sarah Whittle[2,4]

E-mail: swhittle@unimelb.edu.au; Ph: +61 402 597 590

Nicholas B. Allen[5]

E-mail: nallen3@uoregon.edu; Ph: +1 541 346 4075

Karl J. Friston[6] (senior author)

E-mail: k.friston@ucl.ac.uk; Ph: +44 20 3456 7890

[1] Centre for Youth Mental Health, The University of Melbourne, Melbourne, Australia, 3052

[2] Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia, 3010

[3] Orygen, The National Centre of Excellence in Youth Mental Health, Melbourne, Australia, 3052

[4] Melbourne Neuropsychiatry Centre, Department of Psychiatry, The University of Melbourne and Melbourne Health, Melbourne, Australia, 3053

[5] Department of Psychology, University of Oregon, Eugene, Oregon, USA, 97403

[6] Wellcome Trust Centre for Neuroimaging, University College London, London, UK, WC1N3BG

*Correspondence: pbadcock@unimelb.edu.au (P.B. Badcock)

ABSTRACT

Major depressive disorder is a debilitating condition characterised by diverse neurocognitive and behavioural deficits. Nevertheless, our species-typical capacity for depressed mood implies that it serves an adaptive function. Here, we apply an interdisciplinary theory of brain function to explain depressed mood and its clinical manifestations. Combining insights from the free-energy principle with evolutionary theorising in psychology, we argue that depression reflects an adaptive response to perceived threats of aversive social outcomes (e.g., exclusion) that minimises the likelihood of surprising interpersonal exchanges (i.e., those with unpredictable outcomes). We suggest that psychopathology typically arises from ineffectual attempts to alleviate interpersonal difficulties and/or hyper-reactive neurobiological responses to social stress (i.e., uncertainty), which often stems from early experience that social uncertainty is difficult to resolve.

Keywords: Active Inference; Evolutionary Systems Theory; Depression; Free-Energy Principle; Major Depressive Disorder

An Evolutionary Systems Approach to Depression

Why do we become depressed? Why are some of us particularly prone to depression? And how is this best managed? To answer these questions, we require an interdisciplinary approach that synthesises studies of the depressed brain with psychological research on its ecological, ontogenetic and biobehavioural correlates [1, 2]. To this end, we apply an integrative evolutionary systems theory (EST) of human brain function to explain depressed mood and its clinical manifestations. The EST in question rests on two uncontroversial assumptions. The first appeals to a consensus among cognitive scientists that the brain is a hierarchical, self-organising system sculpted by evolution [3-5]. This hierarchy ranges from lower-order, highly specialised neural subsystems responsible for sensory-motor processing; through to highly integrated cortical regions that develop more gradually and underlie the sophisticated, executive cognitive faculties unique to humans [see Box 1]. This calls for a theory of global brain function that explains how depression emerges from coordinated interactions within hierarchically integrated neuronal systems. The second assumption echoes dynamic systems approaches that situate the brain within the evolutionary dynamics of the brain-body-environment system [6-8]. According to this view, the neural mechanisms responsible for depression can only be understood by considering the broader context of human evolution, enculturation, development, embodiment and behaviour.

We aim to exemplify this approach by offering an interdisciplinary hypothesis of the depressed brain. Following the free-energy principle (FEP; see [5]), we first discuss how depressive disorders emerge from the functioning of, and disruptions to, hierarchical neural dynamics that seek to minimise uncertainty. We then integrate this work with psychological research on the adaptive function of depression, along with the familial, developmental and psychobiological mechanisms that often underlie it. We propose that our species-typical capacity for depressed mood can be explained as an evolved biobehavioural strategy that responds adaptively to adverse interpersonal conditions by minimising the likelihood of unpredictable social interactions. We discuss how our model builds on theories of clinical depression in the active inference literature, before turning to the hierarchical neural mechanics that underlie depressed mood and depressive disorder.

Applying the Free-Energy Principle to Depression

The FEP is a global theory of neural structure and function suggesting the brain can be seen as a "prediction machine" that attempts to maximise the evidence for a creature's model of the world by minimising an upper limit on surprise [i.e., free-energy; see Box 2]. In line with predictive coding, the FEP describes the brain as a hierarchical generative model – a hierarchy of hypotheses about the world that enables a reduction of surprise by minimising discrepancies between incoming sensory inputs and top-down predictions [9]. Conditional expectations are thought to be encoded by deep pyramidal cells (i.e., representation units) at each level of the cortical hierarchy that convey predictions downward to suppress errors at the level below, while prediction errors are encoded by superficial pyramidal cells (i.e., error units) that convey errors forward to revise expectations at the level above [10]. This allows us to minimise surprise by updating our internal models (i.e., perception). Alternatively, we can selectively sample sensory data to ensure that our predictions are self-fulfilling – by changing how we act upon the world to confirm our expectations (i.e., active inference [11]). Thus, perception and action operate synergistically to minimise prediction errors and optimise our internal representations of the environment. A key corollary of this model is that our predictions are optimised by evolution, development and learning. Emphasis is placed on adaptive priors – inherited expectations about the way our world unfolds that have been shaped by natural selection to guide action-perception cycles toward adaptive (i.e., unsurprising) states [5, 12].

To date, applications of the FEP to depressive disorders have chiefly concentrated on two processes, stemming from different levels of the cortical hierarchy. The first relates to limbic deficits in minimising prediction error. Barrett and colleagues suggest that depressive disorders arise from aberrant interoceptive predictions originating from abnormalities within the (limbic) agranular visceromotor cortex, which is central to emotional processing, energy regulation and allostatic responses to stress [13, 14]. These abnormalities can arise from past exposure to sustained distress, and generate false (interoceptive) predictions about the body's upcoming autonomic, metabolic and immunological needs that chronically activate physiological stress responses (e.g., dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis and pro-inflammatory states). Over time, the visceromotor systems try to minimise prediction error by producing sickness behaviours (e.g., negative affect and fatigue; also see [15]) that reduce energy expenditure and ultimately manifest in depression [13]. Here, depression is seen as a disorder of allostasis, characterised by energy dysregulation and deficits in interoceptive inference; i.e., an insensitivity to prediction errors and/or a miscalibration of their precision – see glossary [14]. These deficits lead to a failure to update dysfunctional internal models (e.g., cognitive rigidity), perpetuating further metabolic inefficiencies and engendering the downward spiral that typifies depressive illness.

The second class of models concentrates on impairments in top-down expectations of reward. Checkroud [16] has described the depressed brain as a hierarchical constellation of depressive beliefs, which impose a consistent negative bias in predictions that manifests in anhedonic features and the down-regulation of neural reward systems (e.g., dopaminergic and serotonergic dysfunction). In line with active inference, these beliefs exacerbate depression by prompting the individual to actively sample the environment to confirm negative predictions (e.g., learned helplessness). Others have suggested that depressive disorders impair reward-approach behaviours by causing a pathological underconfidence in one's predictions [17], or by distorting higher-order evaluations of the self (e.g., low self-worth), disrupting social behaviour by overweighting the likelihood of aversive interactions [18]. Each of these proposals echo models of reinforcement learning in computational psychiatry and evolutionary biology suggesting that depression emerges from successive discrepancies between actual and expected reward outcomes (i.e., prediction errors), entrenching (empirical) prior beliefs that rewards are unlikely which inhibit reward-approach behaviours [19, 20].

Taken together, the frameworks considered here suggest that depression entails impairments in reward-approach systems emerging from two neurocognitive processes – deficits in the predictive processing of sensory evidence; and prior beliefs that negatively bias predictions. Although this perspective of depression as false inference offers a cohesive, neurobiologically plausible account of the biobehavioural deficits observed in depressive illness, two important questions remain. First, by concentrating on depressive disorders, the models above say little about our species-typical capacity for depressed or low mood. The notable exception is a formal (computational) scheme that defines emotional valence in terms of the rate of change in free-energy over time, with positive and negative affect tracking a decrease and increase in free-energy, respectively (see [17]). In this model, negative moods enable an organism to respond adaptively to unexpected changes in the world by increasing the (learning) rate of evidence accumulation – overweighing recent sensory inputs over prior experiences to heighten sensitivity to environmental change, thereby minimising prediction error [17]. However, this does not specifically address the adaptive significance of depression per se. Second, the literature to date sheds little light on the ecological conditions responsible for the positive selective pressure that depression appears to have [see Box 3]. If depression instantiates an adaptive prior, it should minimise surprise in response to specific environmental challenges that have threatened our inclusive fitness (i.e., free-energy) over evolutionary time. Identifying this adaptive function is arguably central to understanding why depression occurs. To address these issues, we hope to build upon the active inference literature by incorporating complementary insights drawn from an evolutionary systems approach to psychology.

Insights from an Evolutionary Systems Approach to Psychology

In psychology, evolutionary systems models have typically focused on the dynamic interplay between evolutionary and developmental processes (e.g., [8, 21-25]), an approach that has been further extended to reconcile theoretical divisions between major paradigms in the field (see [4]). According to this perspective, the embodied brain and behaviour emerge from selection acting on dynamic interactions between the levels of causation identified by Tinbergen – adaptation, phylogeny, ontogeny, and mechanism [26]. This causal hierarchy is arguably recapitulated by paradigms in psychology, which concentrate differentially on four overlapping levels of explanation – ultimate hypotheses for adaptive, species-typical characteristics (i.e., evolutionary psychology); epigenetic explanations for intergenerational, between-group differences (i.e., evolutionary developmental biology and psychology); dynamical explanations for individual similarities and differences (i.e., developmental psychology); and mechanistic explanations for real-time phenomena (i.e., psychological subdisciplines such as cognitive, biological, personality, social and clinical psychology) [4]. Central to EST is the need to explore how these causal levels interact – evolutionary influences on neural structure and function constrain individual development and learning, while effects at these more proximate levels can shape the evolution of the brain [3, 27]. To explain depression then, we require a multi-level hypothesis that synthesises diverse fields of psychological inquiry to explain both why it is adaptive, along with how it emerges from intergenerational, developmental and real-time mechanisms.

Although there are various Darwinian models of depression [28], a theme common to many of these is that low mood reflects an adaptive biobehavioural strategy that conserves or reallocates energy and resources in unpropitious social environments [29, 30]. According to this view, depressed mood states are elicited by aversive interpersonal outcomes (e.g., exclusion, defeat, or loss) that indicate a critical loss of control over social relationships that were critical to ancestral fitness [31]. A model that incorporates influential theories in this area and shows promising conceptual parallels with the FEP is the social risk hypothesis (technically, risk corresponds to uncertainty and uncertainty is expected surprise or free energy). This maintains that depressed mood reflects an adaptive, risk-averse approach to social interaction that reduces the likelihood of further aversive outcomes by: (1) increasing our cognitive sensitivity to (sensory) cues of social risk; (2) reducing our (behavioural) propensity for taking social risks; and (3) initiating signalling behaviours that elicit support and defuse conflict [32].

The idea that depression reflects an evolved response to adverse social conditions concords with evidence that extends across Tinbergen's remaining levels of inquiry. The intergenerational transmission of susceptibility to depressive disorders due to deleterious social environments is widely documented [33, 34], with studies involving rodents, primates, and humans showing that exposure to social stressors (e.g., low maternal care and social defeat) produces potentially heritable epigenetic changes that confer risk for disorder by heightening stress reactivity [35, 36]. Ontogenetically, exposure to early social stress (e.g., parental loss, abuse, or neglect) is a strong predictor of depressive vulnerability [37], and is thought to heighten susceptibility to disorder by leading to hyperactivity of the HPA axis [38, 39] and up-regulating pro-inflammatory immune responses [40]. Behavioural and neuroimaging studies further suggest that the risk of depressive onset rises markedly in adolescence because of an increased sensitivity to social threats in this period [41, 42]. Finally, research across the sub-disciplines highlights an intimate connection between depression and the social world (see [43]). For example, the precipitants of depression are typically interpersonal in nature [44]; social support and belonging are key protective factors [45]; and typical correlates of depression clearly exemplify negative self-other relations (e.g., low self-esteem [46]). Consistent with the social risk hypothesis, there are also multiple lines of evidence to suggest that low mood is associated with biobehavioural changes that facilitate adaptive responses to social stress. Depressive cognition is characterised by a specific, attentional bias towards socially-threatening stimuli [47] and increased rumination about interpersonal problems [48], while normative depressed states have been shown to increase the accuracy of social inferences (e.g., depressive realism [49]) and improve social problem-solving [50]. Furthermore, many features of depression – such as anhedonia, a negative thinking bias and social withdrawal – reduce exposure to social risks by inhibiting reward-approach behaviours [51], while the signalling behaviours associated with depression (e.g., reassurance-seeking and submissive behaviours) explicitly attempt to elicit support and defuse potential conflict [52-54]. Notably, other studies have provided direct empirical support for the social risk hypothesis itself (Badcock & Allen, 2003; Badcock & Allen, 2007; Dunn, Whelton & Sharpe, 2012; Girard, Cohn, Mahoor, Mavadati, & Rosenwald, 2013).

In light of such work, we suggest that the human capacity for depressed mood can be explained in terms of a risk-averse adaptive prior that minimises uncertainty in the social world when sensory cues indicate both a high degree of socio-environmental volatility (i.e., unpredictability) and an increased probability of aversive interpersonal outcomes (e.g., rejection, defeat or loss) (see Figure 1). This depressive response instantiates a "better safe than sorry" strategy that minimises the likelihood of unpredictable social interactions by causing adaptive changes in cognition (e.g., hypersensitivity to aversive social stimuli, a negative thinking bias and deficits in responses to reward) and action (e.g., risk-averse behaviours such as social withdrawal). Epigenetic and ontogenetic mechanisms arguably support this function by sensitising the individual to socio-environmental volatility when developmental insults indicate that the probability of aversive social interactions is high, producing hyper-reactive stress response systems that heighten risk for disorder by increasing the precision of social prediction errors and prompting exaggerated, pathological responses to interpersonal stressors.

Notably, the exacerbation of normative depressed states into severe, dysfunctional forms is also likely when depressive changes fail to alleviate social stress, creating a self-perpetuating cycle arising from heightened and prolonged arousal of ineffectual attempts to reduce socio-environmental volatility [32]. This, in turn, is likely to chronically activate neurophysiological stress responses and leads to debilitating sickness behaviours [13, 14]. As discussed above, previous applications of the FEP suggest that this depressive spiral is engendered by a positive feedback loop between two neurocognitive mechanisms – the increased precision of social prediction errors, coupled with a negative bias affecting social predictions. Following active inference, this is likely to engender ongoing depressive behaviours that seek to confirm negative biases, creating a self-fulfilling prophecy (i.e., high predictability) born from mutually reinforcing patterns of cognition and behaviour [16]. Here, depressed can be interpreted as a maladaptive pattern of dysregulated defences – if this depressive response is effective, an individual either escapes or avoids the social stressor or adapts to it; if the defence fails, the individual is at risk of entering a self-perpetuating dysregulated state, which falls beyond the normal range of adaptive functioning (Allen & Badcock, 2003; Gilbert, 2001). Nevertheless, it should also be recognised that clinical manifestations of depression can result from asocial causes that produce neurobiological abnormalities typically associated with dysregulated mood (e.g., pro-inflammatory immune responses induced by illness and medications; 40).

The Depressed Brain

It is widely accepted that depression emerges from bidirectional interactions between hierarchically organised neural regions. Most of the theoretical work in this area concentrates on two general brain systems that work in concert – a ventral affective system, including subcortical regions such as the amygdala and ventral striatum; and the prefrontal cortex (PFC), which modulates the reactions of the ventral affective system [1]. These systems are composed of subcortical neural circuits responsible for the unconscious processing of affective and social stimuli on the one hand; and on the other, executive networks that regulate affective states, with medial prefrontal regions playing a particularly important role in modulating visceral and behavioural responses in order to adapt them to the external milieu [1].

More particularly, evidence gleaned from neuroimaging and animal studies suggests that depression involves dysfunction of the "extended visceromotor network", in which the medial PFC regulates affective states by modulating visceromotor output via connections with the amygdala, ventral striatum, hypothalamus, and other subcortical regions [55]. Brain regions across this network regulate motivation (e.g., anhedonia and dopaminergic function) and neurobiological responses to stress, and play a central role in social threat and reward processing [39, 41, 56, 57]. Neurodevelopmental changes in these regions throughout adolescence are also thought to heighten vulnerability to disorder by increasing sensitivity to rapidly changing social contexts in this period (see Box 4). Collectively, such findings fit well with our proposal that depression often stems from the need to adapt to complex social contexts, and manifests through the bidirectional interplay of hierarchical neuronal processes.

Specifically, we speculate that the extended visceromotor system responds to volatility in the social environment by increasing the precision of social prediction errors, initiating changes in neuronally encoded expectations that increase attention to social cues and motivate risk-averse behaviours (e.g., social withdrawal). This heightened sensitivity to somatic and affective cues leads, in turn, to further avoidance of interpersonal stressors. The depressive response is adaptive when changes in mood state and behaviour reduce uncertainty in the face of socio-environmental change, and lead to re-engagement with that environment when volatility abates (which should, at least in part, be brought about by depressive behaviours; see [32]). However, following the active inference literature, we suggest that the depressive response becomes maladaptive when there are (neuromodulatory) failures of “precision engineered” visceromotor inference – produced, for instance, by sustained social distress – leading to illness behaviours which fail to respond to improvements in interpersonal contexts and can often exacerbate socio-environmental stress [13, 14]. Neurodevelopmentally, the PFC can also potentiate vulnerability to depression by underwriting the formation of distal goals that, when frustrated by rejection or failure, can lead to depression by suppressing the brain’s reward system [58] and the confidence in (or precision of) our beliefs about behaviour [59], thereby inhibiting goal-directed behaviours.

Ultimately, our basic claim is that depression can be viewed as an adaptive faculty that underwrites emotional allostasis in an increasingly prosocial and volatile world. Physiologically, this faculty increases sensitivity to interpersonal, affiliative and interoceptive cues. Clearly, sensitisation to stressful exteroceptive and interoceptive cues also has to be predicted by the hierarchical brain, which implicates the functional neuroanatomy described above. Under active inference, sensitivity to stress-related cues corresponds to their precision [13, 60], implicating neuromodulatory systems associated with reward, action selection and interoceptive inference [61-63]. Crucially, in order to act it is also necessary to attenuate the precision afforded to the sensory consequences of action (i.e., we have to ignore the fact that we are not currently acting). This means that an adaptive depressive response suspends sensory attenuation – and action – so that we can attend to interpersonal prediction errors and revise our (posterior) beliefs about our relationships with others, via perceptual inference and learning. Sensory attenuation can be regarded as the complement of sensory attention; i.e., attenuating or augmenting the gain (precision) afforded sensory prediction errors to ignore or select sensory information, respectively. According to this scheme, maladaptive forms of depression reflect a pervasive, self-maintaining failure of sensory attenuation, leading to ruminations, false inference and a concomitant inability to act and test these false beliefs.

Interestingly, exactly the same conclusions (namely, a failure of sensory attenuation) have been drawn for a range of neuropsychiatric disorders, ranging from autism [64] to schizophrenia [65]. One could ask what is specific about this mechanism in depression, and respond by referring to the particular (interoceptive and affiliative) modalities affected. However, perhaps the more intriguing implication is that the comorbidity of depression and other disorders might arise from a common pathophysiological mechanism, which can be explained in terms of false inference.

Concluding Remarks

In this opinion piece, we have endeavoured to contribute to the active inference literature on mood disorder by suggesting that normative levels of depressed mood instantiate an adaptive prior that minimises the likelihood of surprising interpersonal interactions when faced with threats of aversive social outcomes that typically compromised ancestral fitness. By extending beyond previous applications of the FEP to emphasise both the adaptive function of low mood and the causal role of the social ecology, we believe our model demonstrates the heuristic benefits of combining active inference with insights in psychology to improve our understanding of depressed mood and mood disorder. It also motivates new questions for research, calling for greater integration between neuroscientific and psychological approaches to explore the ways in which the neural mechanisms that underpin depression relate to behaviour, development and the social world [see Outstanding Questions]. In particular, the idea that depression can emerge from the need to navigate social risks stands to inform theory-driven approaches in computational psychiatry, which improve our understanding, prediction and treatment of mental illness by using simulations and mathematical models to capture complex interactions across multiple causal levels [19, 66].

That said, we do not wish to imply that depression is solely attributable to social causes. In evolutionary psychology, for instance, the distinction between social and non-social depressive responses is widely recognised [Durisko 2015; Gilbert 2006], and as we have noted, depression can also arise from depressogenic neuroanatomical abnormalities produced by influences other than unfavourable social conditions. Nevertheless, our model adds to the active inference literature by emphasising the importance of the social environment in explaining the aetiology and phenomenology of depression. This underscores the need to develop (computational) diagnostic tools that are capable of distinguishing between social and non-social forms in order to inform treatment decisions.

In closing, our model also promotes clear avenues for intervention. To date, proponents of the FEP have advocated treatments that directly target dysregulated neural systems, such as psychopharmacological agents that act upon the neurotransmitter systems that encode precision or uncertainty (e.g., serotonin and dopamine [16]). They have also recommended the use of cognitive behavioural therapies to disrupt the spiral of self-defeating actions typical of depression [16], or to construct new prediction signals that modify the gain on prediction errors via the salience network [14]. Our own model adds to this work by emphasising the need to facilitate adaptive responses to social stress. This could well explain the efficacy of interpersonal psychotherapy as a treatment for major depressive disorder [67], and highlights the value of prevention and early intervention efforts that reduce vulnerability by targeting modifiable risk factors in the social environment. Given the heterogeneous nature of depression, we also recommend the development of (computational) diagnostic tools capable of distinguishing between social and non-social forms in order to inform treatment decisions. Finally, simply having a positive and principled framework within which to understand depression – and the rationale for therapeutic interventions – is likely to be helpful for those seeking treatment. Our synthesis can be used to help clients understand why they have depression, and to explain why, for example, it might be useful to combine interpersonal psychotherapy with antidepressants.

Acknowledgements: We would like to thank Lucy Morrish, Jakob Hohwy, Alex Fornito, Rebecca Schwarzlose and three anonymous reviewers for their valuable contributions. Karl Friston is funded by the Wellcome Trust; Christopher Davey and Sarah Whittle are both funded by the National Health and Medical Research Council of Australia (NHMRC).

Glossary

Active inference: A corollary of the free-energy principle, which states that we minimise surprise (i.e., prediction errors) by changing our predictions (i.e., perception) or by acting upon the world to elicit sensations that conform to predictions (i.e., action).

Adaptive prior: A prior endowed by evolution to underwrite adaptive fitness.

Association cortex: Regions of the cerebral cortex that are not primary sensory or motor projection areas, including the prefrontal cortex, and extensive parts of the temporal, parietal and occipital cortices..

Empirical priors: Priors found in hierarchical models that can be learned or inferred under priors from the level above.

Entropy: The uncertainty or average surprise associated with outcomes sampled from a probability distribution. A distribution with low entropy means, on average, that the outcome is relatively predictable.

Evolutionary systems theory: A multidisciplinary paradigm that explains dynamic, evolving systems in terms of co-action between self-organisation and general selection (e.g., natural selection) over time. This produces complex adaptive systems, like the brain, that adapt to the environment through an autonomous process of selection that recruits the outcomes of locally interacting components within that system to select a subset of those components for replication or enhancement.

Free-energy principle: A generalisation of predictive coding that asserts that organisms actively minimise an upper bound on surprise (i.e., free energy), which, under simplifying assumptions, translates to (precision weighted) prediction error.

Generative Model: A probabilistic mapping from hidden causes in the environment to observed consequences (sensory data), typically specified in terms of the likelihood of observing some data (given their causes) and priors (on these causes).

Interoception: The perception and integration of autonomic, hormonal, visceral and immunological (bodily) signals.

Precision: The inverse variance, volatility, or reliability of a signal. In predictive coding, prediction errors are weighted by precisions that determine the relative influence of bottom-up (error) and top-down (representation) signals (e.g., a high precision on error signals corresponds to low confidence in top-down beliefs). Dynamic precision weighting is mediated by neuromodulation and underwrites psychological processes such as attentional selection and sensory attenuation.

Predictive coding: A processing scheme for inferring the likely causes of sensory data by minimising prediction error. Typically, this entails a hierarchical generative model (e.g., the brain) in which top-down signals convey predictions and bottom-up signals convey (precision weighted) prediction errors.

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

Surprise: The negative log probability of sensory experiences encountered by an agent. Also known as surprisal or self-information.

Visceromotor cortex: Agranular (limbic) regions of isocortex and allocortex that regulate the hormonal, immune and autonomic nervous systems, including the cingulate cortex, the posterior ventral medial prefrontal cortex, the posterior orbitofrontal cortex and ventral portions of the anterior insula.

BOX 1: The Hierarchical Structure of the Brain

In psychology, it has long been recognised that the brain entails a hierarchical structure ranging from highly specialised sensorimotor systems at its lowest levels through to developmentally flexible, highly integrated systems responsible for higher-order executive functions [3, 4]. A hierarchical neural architecture is also emphasised by predictive coding approaches in neuroscience, which explore how the brain minimises prediction error via recurrent message-passing between cortical levels [9, 68, 69]. More recently, imaging studies in network neuroscience have provided direct evidence that the brain exhibits a multiscale hierarchical organisation, with a given node (e.g., network, module or sub-module) itself comprising a network of smaller interacting nodes at a lower level [68, 70] (see Figure I).

Comparative work suggests that a hierarchical architecture is a hallmark of the mammalian brain, progressing from highly segregated sensorimotor hierarchies common to all mammals through to the cortical association areas that confer the adaptive advantage of heightened cognitive control among primates [71, 72]. Again, this structure is thought to exemplify the complementary relationship between evolution and development – selection has canalised early sensorimotor regions that serve as neurodevelopmental anchors, allowing for the progressive, activity-dependent self-organisation of widely distributed association networks that lie furthest from sensory patterning centres [71, 73]. This neuroplasticity enhances adaptability by producing higher-order, “domain-general” faculties that are able to respond flexibly to rapidly changing environments [6, 73].

It is now broadly accepted that a hierarchical neural structure is favoured by selection. It enhances evolvability because deleterious changes to a single component of the system are unlikely to affect the system itself, and it allows adaptive novelties to emerge without disrupting global functioning [70]. Computer simulations of evolving networks have also shown that selection favours a hierarchical organisation because it conserves the (spatial, processing and metabolic) cost of neural connections; improves problem-solving by recursively combining solutions to sub-problems; and adapts more rapidly to new environments than non-hierarchical structures [74]. Finally, the hierarchical brain is thought to promote “self-organised criticality”. This is a dynamical state poised between completely ordered, stable cycles of activity and highly complex, chaotic ones that optimises evolvability because it allows small extrinsic changes to elicit large intrinsic reorganisations. The hierarchical segregation of neural networks into distributed neighbourhoods has been found to stretch the parameter range for self-organised criticality by allowing subcritical and supercritical dynamics to co-exist simultaneously [75]. Since systems at criticality have optimal information-processing capacities, a structure that extends this critical region is likely to be naturally selected [76].

BOX 2: The Free-Energy Principle

The FEP seeks to explain how biological systems maintain their integrity by occupying a constricted number of states [5]. It suggests that all organisms actively reduce the entropy (i.e. disorder or dispersion) of their sensory and physical states by minimising free-energy. Borrowed from statistical thermodynamics and machine learning, free-energy is an information theory quantity which limits (by being greater than) the entropy of a brain's sensations or sensory samples from the environment. In this context, entropy (the mathematical description of uncertainty) refers to the (long-term) average of surprise: a statistical concept referring to the negative log probability of sensory samples encountered by an agent. This probability is also known as (Bayesian) model evidence.

These principles have important implications for understanding how biotic agents work. Because the repertoire of states an organism occupies is limited, the probability of these states has low entropy (i.e., surprise). Thus, an organism's distal imperative of maintaining functional states within physiological bounds (i.e., homeostasis) translates into a proximal avoidance of surprise [5]. Surprise itself cannot be evaluated; however, biological systems can minimise surprise vicariously by minimising their free-energy – which roughly translates to prediction error, weighted by its precision.

The FEP appeals to predictive coding by characterising the brain as a hierarchical inference machine that minimises prediction error by seeking to match incoming sensory inputs with top-down predictions [see Figure II]. This occurs in two ways. First, we can improve our predictions by altering internal states (i.e., perception). Second, we can act upon the world to confirm our predictions (i.e., action). Thus, action and perception operate synergistically to optimise an organism's model of the environment. Crucially, to minimise free-energy, the precision of prediction errors also has to be predicted, invoking notions of attentional gain (psychologically) and neuromodulation (physiologically).

The FEP also applies to the morphology, development and evolution of the brain. It suggests that instead of just containing a model of the world, the brain is a model of the world – a physical transcription of causal regularities in the environment that is optimised by evolution. This model instantiates genetically specified (empirical) prior beliefs that have minimised free-energy (i.e., maximised model evidence) over evolutionary time by ensuring an organism seeks out a small number of unsurprising states that are consistent with its phenotype and environment. In other words, natural selection is nature's way of performing Bayesian model selection to minimise the (variational) free energy of phenotypes (i.e., generative models).

Box 3: The Adaptive Significance of Depression

Box 4: The Adolescent Brain and Risk for Depression

The brain undergoes significant maturation in adolescence, involving processes that begin with puberty and continue until a young person is in their mid-to-late twenties [77]. Over this period, there is a progressive increase in white matter, alongside synaptic pruning and grey matter loss, which have the effect of delineating more clearly defined large-scale brain networks [78]. Subcortical regions, including the primary components of the reward system, undergo more rapid maturation [79], while the most prolonged development is in association cortex, including prefrontal regions that are implicated in social processing [78, 80].

It is now widely accepted that the functional and structural changes that accompany adolescence reflect a particularly sensitive period for adapting to the social world. Brain imaging studies show that adolescence is typified by significant alterations in social and affective processing systems, which are thought to increase risk for mood disorder by heightening sensitivity to social threats in this period [41, 80-82]. Coincident with these neurodevelopmental processes, there are also substantial changes in the adolescent social environment. Peer relationships become increasingly important, hierarchical and complex, and there is significant socio-environmental volatility – friendships change frequently, and romantic relationships are typically short-lived [83].

It is unsurprising, then, that the period from adolescence to early adulthood is a peak time for the onset of depression [42]. During adolescence, sources of social uncertainty are frequently encountered. Maturation of subcortical regions, along with marked hormonal changes [82, 84], increase sensitivity to affective and self-relevant social cues. Moreover, prefrontal cortical development leads, on the one hand, to improved regulation of affective processes, but on the other, heightens sensitivity to the nuance and complexity of interpersonal relationships [58]. For this reason, increased vulnerability to depression starts in puberty but is maintained well beyond adolescence.

References

1. Pfeifer, J.H. and Allen, N.B. (2012) Arrested development? Reconsidering dual-systems models of brain function in adolescence and disorders. Trends Cognit. Sci. 16, 322-329.
2. Pfeifer, J.H. and Allen, N.B. (2016) The audacity of specificity: Moving adolescent developmental neuroscience towards more powerful scientific paradigms and translatable models. Dev. Cogn. Neurosci. 17, 131-137.
3. Badcock, P.B. (2012) Evolutionary systems theory: A unifying meta-theory of psychological science. Rev. Gen. Psychol. 16, 10.
4. Friston, K. (2010) The free-energy principle: A unified brain theory? Nat. Rev. Neurosci. 11, 127-138.
5. Badcock, P. et al. (submitted) The hierarchically mechanistic mind: An evolutionary systems theory of the brain and behaviour.
6. Anderson, M.L. (2014) *After Phrenology*, MIT Press.
7. Clark, A. (2015) *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Oxford University Press.
8. Lickliter, R. and Honeycutt, H. (2003) Developmental dynamics: toward a biologically plausible evolutionary psychology. Psychol. Bull. 129, 819-835.
9. Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain. Sci. 36, 181-204.
10. Bastos, A.M. et al. (2012) Canonical microcircuits for predictive coding. Neuron 76, 695-711.
11. Friston, K.J. et al. (2010) Action and behavior: A free-energy formulation. Biol. Cybern. 102, 227-260.
12. Friston, K. (2013) Life as we know it. J. R. Soc. Interface 10, 20130475.

13. Barrett, L.F. and Simmons, W.K. (2015) Interoceptive predictions in the brain. Nat. Rev. Neurosci. 16, 419-429.

14. Barrett, L.F. et al. (2016) An active inference theory of allostasis and interoception in depression. Philos. Trans. R. Soc. Lond. B Biol. Sci. 371, 20160011.

15. Seth, A.K. and Friston, K.J. (2016) Active interoceptive inference and the emotional brain. Philos. Trans. R. Soc. Lond. B Biol. Sci. 371, 20160007.

16. Chekroud, A.M. (2015) Unifying treatments for depression: An application of the Free Energy Principle. Front Psychol 6, 153.

17. Joffily, M. and Coricelli, G. (2013) Emotional valence and the free-energy principle. PLoS Comput. Biol. 9, e1003094.

18. Moutoussis, M. et al. (2014) Bayesian inferences about the self (and others): A review. Conscious. Cogn. 25, 67-76.

19. Adams, R.A. et al. (2016) Computational psychiatry: Towards a mathematically informed understanding of mental illness. J. Neurol. Neurosurg. Psychiatry 87, 53-63.

20. Nettle, D. and Bateson, M. (2012) The evolutionary origins of mood and its disorders. Current Biology 22, R712-R721.

21. Caporael, L.R. (2001) Evolutionary psychology: Toward a unifying theory and a hybrid science. Annu. Rev. Psychol. 52, 607-628.

22. Geary, D.C. and Bjorklund, D.F. (2000) Evolutionary developmental psychology. Child Dev. 71, 57-65.

23. Kenrick, D.T. et al. (2002) Dynamical evolutionary psychology: Mapping the domains of the new interactionist paradigm. Pers. Soc. Psychol. Rev. 6, 347-356.

24. Ploeger, A. et al. (2008) Is evolutionary psychology a metatheory for psychology? A discussion of four major issues in psychology from an evolutionary developmental perspective. Psychol. Inq. 19, 1-18.

25. Frankenhuis, W.E. et al. (2013) Bridging developmental systems theory and evolutionary psychology using dynamic optimization. Dev. Sci. 16, 584-598.

26. Tinbergen, N. (1963) On aims and methods of ethology. Z. Tierpsychol. 20, 410-433.

27. Marshall, P.J. (2013) Coping with complexity: Developmental systems and multilevel analyses in developmental psychopathology. Dev. Psychopathol. 25, 1311-1324.

28. Durisko, Z. et al. (2015) An adaptationist perspective on the etiology of depression. J. Affect. Disord. 172, 315-323.

29. Beck, A.T. and Bredemeier, K. (2016) A unified model of depression integrating clinical, cognitive, biological, and evolutionary perspectives. Clin. Psychol. Sci. 4, 596-619.

30. Trimmer, P.C. et al. (2015) Adaptive learning can result in a failure to profit from good conditions: Implications for understanding depression. Evol. Med. Public Health 2015, 123-135.

31. Gilbert, P. (2006) Evolution and depression: Issues and implications. Psychol. Med. 36, 287-297.

32. Allen, N.B. and Badcock, P.B. (2003) The social risk hypothesis of depressed mood: Evolutionary, psychosocial, and neurobiological perspectives. Psychol. Bull. 129, 887-913.

33. Vialou, V. et al. (2013) Epigenetic mechanisms of depression and antidepressants action. Annu. Rev. Pharmacol. Toxicol. 53, 59-87.

34. Weissman, M.M. et al. (2005) Families at high and low risk for depression: A 3-generation study. Arch. Gen. Psychiatry 62, 29-36.

35. Sun, H. et al. (2013) Epigenetics of the depressed brain: Role of histone acetylation and methylation. Neuropsychopharmacol. 38, 124-137.

36. Meaney, M.J. (2001) Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. Annu. Rev. Neurosci. 24, 1161-1192.

37. Heim, C. and Binder, E.B. (2012) Current research trends in early life stress and depression: Review of human studies on sensitive periods, gene–environment interactions, and epigenetics. Exp. Neurol. 233, 102-111.

38. Gold, P. (2015) The organization of the stress system and its dysregulation in depressive illness. Mol. Psychiatry 20, 32-47.

39. De Raedt, R. and Koster, E.H. (2010) Understanding vulnerability for depression from a cognitive neuroscience perspective: A reappraisal of attentional factors and a new conceptual framework. Cogn. Affect. Behav. Neurosci. 10, 50-70.

40. Slavich, G.M. and Irwin, M.R. (2014) From stress to inflammation and major depressive disorder: A social signal transduction theory of depression. Psychol. Bull. 140, 774.

41. Silk, J.S. et al. (2012) Why do anxious children become depressed teenagers? The role of social evaluative threat and reward processing. Psychol. Med. 42, 2095-2107.

42. Andersen, S.L. and Teicher, M.H. (2008) Stress, sensitive periods and maturational events in adolescent depression. Trends Neurosci. 31, 183-191.

43. Gotlib, I.H. and Hammen, C. (2014) *Handbook of Depression* (3rd edn.), Guilford Press.

44. Vrshek-Schallhorn, S. et al. (2014) Refining the candidate environment interpersonal stress, the serotonin transporter polymorphism, and gene-environment interactions in major depression. Clin. Psychol. Sci. 2, 235-248.

45. Cohen, S. and Wills, T.A. (1985) Stress, social support, and the buffering hypothesis. Psychol. Bull. 98, 310-357.

46. Hawkley, L.C. and Capitanio, J.P. (2015) Perceived social isolation, evolutionary fitness and health outcomes: A lifespan approach. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370, 20140114.

47. Leppänen, J.M. (2006) Emotional information processing in mood disorders: A review of behavioral and neuroimaging findings. Curr. Opin. Psychiatry 19, 34-39.

48. Hankin, B.L. et al. (2010) Corumination, interpersonal stress generation, and internalizing symptoms: Accumulating effects and transactional influences in a multiwave study of adolescents. Dev. Psychopathol. 22, 217-235.

49. Moore, M.T. and Fresco, D.M. (2012) Depressive realism: A meta-analytic review. Clin. Psychol. Rev. 32, 496-509.

50. Forgas, J.P. (2016) Can sadness be good for you? Aust. Psychol. DOI: 10.1111/ap.12230

51. Pizzagalli, D.A. (2014) Depression, stress, and anhedonia: Toward a synthesis and integrated model. Annu. Rev. Clin. Psychol. 10, 393.

52. Hagen, E.H. (2011) Evolutionary theories of depression: A critical review. Can. J. Psychiatry 56, 716-726.

53. Hames, J.L. et al. (2013) Interpersonal processes in depression. Annu. Rev. Clin. Psychol. 9, 355-377.

54. Sloman, L. and Gilbert, P. (2000) *Subordination and Defeat: An Evolutionary approach to Mood Disorders and Their Therapy*, Routledge.

55. Price, J.L. and Drevets, W.C. (2012) Neural circuits underlying the pathophysiology of mood disorders. Trends Cogn. Sci. 16, 61-71.

56. Kupferberg, A. et al. (2016) Social functioning in major depressive disorder. Neurosci. Biobehav. Rev. 69, 313-332.

57. Rushworth, M.F. et al. (2013) Are there specialized circuits for social cognition and are they unique to humans? Curr. Opin. Neurobiol. 23, 436-442.

58. Davey, C.G. et al. (2008) The emergence of depression in adolescence: Development of the prefrontal cortex and the representation of reward. Neurosci. Biobehav. Rev. 32, 1-19.

59. Friston, K. et al. (2015) Active inference and epistemic value. Cogn. Neurosci., 1-28.

60. Friston, K. et al. (2014) The anatomy of choice: Dopamine and decision-making. Philos. Trans. R. Soc. Lond. B Biol. Sci. 369, 20130481.

61. Paulus, M.P. and Stein, M.B. (2006) An insular view of anxiety. Biol. Psychiatry 60, 383-387.

62. Pezzulo, G. et al. (2015) Active Inference, homeostatic regulation and adaptive behavioural control. Prog. Neurobiol. 134, 17-35.

63. Waselus, M. et al. (2011) Collateralized dorsal raphe nucleus projections: A mechanism for the integration of diverse functions during stress. J. Chem. Neuroanat. 41, 266-280.

64. Pellicano, E. and Burr, D. (2012) When the world becomes ‘too real’: A Bayesian explanation of autistic perception. Trends Cogn. Sci. 16, 504-510.

65. Fletcher, P.C. and Frith, C.D. (2009) Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. Nat. Rev. Neurosci. 10, 48-58.

66. Friston, K.J. et al. (2014) Computational psychiatry: The brain as a phantastic organ. Lancet Psychiatry 1, 148-158.

67. Cuijpers, P. et al. (2011) Interpersonal psychotherapy for depression: A meta-analysis. Am. J. Psychiatry 168, 581-592.

68. Park, H.-J. and Friston, K. (2013) Structural and functional brain networks: from connections to cognition. Sci. 342, 1238411.

69. Hohwy, J. (2013) *The Predictive Mind*, Oxford University Press.

70. Sporns, O. and Betzel, R.F. (2016) Modular brain networks. Annu. Rev. Psychol. 67, 613-640.

71. Buckner, R.L. and Krienen, F.M. (2013) The evolution of distributed association networks in the human brain. Trends Cogn. Sci. 17, 648-665.

72. Finlay, B.L. and Uchiyama, R. (2015) Developmental mechanisms channeling cortical evolution. Trends Neurosci. 38, 69-76.

73. Anderson, M.L. and Finlay, B.L. (2014) Allocating structure to function: The strong links between neuroplasticity and natural selection. Front. Hum. Neurosci. 7, 918.

74. Mengistu, H. et al. (2016) The evolutionary origins of hierarchy. PLOS Comput. Biol. 12, e1004829.

75. Hilgetag, C.C. and Hütt, M.-T. (2014) Hierarchical modular brain connectivity is a stretch for criticality. Trends Cogn. Sci. 18, 114-115.

76. Hesse, J. and Gross, T. (2014) Self-organized criticality as a fundamental property of neural systems. Front. Syst. Neurosci. 8, 46-59.

77. Sowell, E.R. et al. (2003) Mapping cortical change across the human life span. Nat. Neurosci. 6, 309-315.

78. Blakemore, S.-J. (2012) Imaging brain development: the adolescent brain. Neuroimage 61, 397-406.

79. Spear, L.P. (2000) The adolescent brain and age-related behavioral manifestations. Neurosci. Biobehav. Rev. 24, 417-463.

80. Fuhrmann, D. et al. (2015) Adolescence as a sensitive period of brain development. Trends Cogn. Sci. 19, 558-566.

81. Blakemore, S.-J. and Mills, K.L. (2014) Is adolescence a sensitive period for sociocultural processing? Annu. Rev. Psychol. 65, 187-207.

82. Crone, E.A. and Dahl, R.E. (2012) Understanding adolescence as a period of social–affective engagement and goal flexibility. Nat. Rev. Neurosci. 13, 636-650.

83. Brown, B.B. and Larson, J. (2009) Peer relationships in adolescence. In Handbook of Adolescent Psychology (3rd edn) (Lerner, R.M. and Steinberg, L. eds), pp. 74-103, John Wiley & Sons.

84. Sisk, C.L. and Zehr, J.L. (2005) Pubertal hormones organize the adolescent brain and behavior. Front. Neuroendocrinol. 26, 163-174.

85.

1 Pfeifer, J.H. and Allen, N.B. (2012) Arrested development? Reconsidering dual-systems models of brain function in adolescence and disorders. *Trends in cognitive sciences* 16, 322-329

2 Pfeifer, J.H. and Allen, N.B. (2016) The audacity of specificity: Moving adolescent developmental neuroscience towards more powerful scientific paradigms and translatable models. *Developmental cognitive neuroscience* 17, 131-137

3 Badcock, P.*, et al.* (submitted) The Hierarchically Mechanistic Mind: An evolutionary systems theory of the brain and behaviour. .

4 Badcock, P.B. (2012) Evolutionary systems theory: A unifying meta-theory of psychological science. *Review of General Psychology* 16, 10

5 Friston, K. (2010) The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience* 11, 127-138

6 Anderson, M.L. (2014) *After phrenology*. MIT Press

7 Clark, A. (2015) *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press

8 Lickliter, R. and Honeycutt, H. (2003) Developmental dynamics: toward a biologically plausible evolutionary psychology. *Psychological bulletin* 129, 819

9 Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* 36, 181-204

10 Bastos, A.M.*, et al.* (2012) Canonical microcircuits for predictive coding. *Neuron* 76, 695-711

11 Friston, K.J.*, et al.* (2010) Action and behavior: a free-energy formulation. *Biological cybernetics* 102, 227-260

12 Friston, K. (2013) Life as we know it. *Journal of the Royal Society Interface* 10, 20130475

13 Barrett, L.F. and Simmons, W.K. (2015) Interoceptive predictions in the brain. *Nature Reviews Neuroscience* 16, 419-429

14 Barrett, L.F.*, et al.* (2016) An active inference theory of allostasis and interoception in depression. *Phil. Trans. R. Soc. B* 371, 20160011

15 Seth, A.K. and Friston, K.J. (2016) Active interoceptive inference and the emotional brain. *Phil. Trans. R. Soc. B* 371, 20160007

16 Chekroud, A.M. (2015) Unifying treatments for depression: an application of the Free Energy Principle. *Frontiers in psychology* 6

17 Joffily, M. and Coricelli, G. (2013) Emotional valence and the free-energy principle. *PLoS Comput Biol* 9, e1003094

18 Moutoussis, M.*, et al.* (2014) Bayesian inferences about the self (and others): A review. *Consciousness and cognition* 25, 67-76

19 Adams, R.A.*, et al.* (2016) Computational Psychiatry: towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery & Psychiatry* 87, 53-63

20 Nettle, D. and Bateson, M. (2012) The evolutionary origins of mood and its disorders. *Current Biology* 22, R712-R721

21 Caporael, L.R. (2001) Evolutionary psychology: Toward a unifying theory and a hybrid science. *Annual review of psychology* 52, 607-628

22 Geary, D.C. and Bjorklund, D.F. (2000) Evolutionary developmental psychology. *Child development* 71, 57-65

23 Kenrick, D.T.*, et al.* (2002) Dynamical evolutionary psychology: Mapping the domains of the new interactionist paradigm. *Personality and Social Psychology Review* 6, 347-356

24 Ploeger, A.*, et al.* (2008) Is evolutionary psychology a metatheory for psychology? A discussion of four major issues in psychology from an evolutionary developmental perspective. *Psychological Inquiry* 19, 1-18

25 Frankenhuis, W.E.*, et al.* (2013) Bridging developmental systems theory and evolutionary psychology using dynamic optimization. *Developmental Science* 16, 584-598

26 Tinbergen, N. (1963) On aims and methods of ethology. *Zeitschrift für Tierpsychologie* 20, 410-433

27 Marshall, P.J. (2013) Coping with complexity: Developmental systems and multilevel analyses in developmental psychopathology. *Development and psychopathology* 25, 1311-1324

28 Durisko, Z.*, et al.* (2015) An adaptationist perspective on the etiology of depression. *Journal of affective disorders* 172, 315-323

29 Beck, A.T. and Bredemeier, K. (2016) A unified model of depression integrating clinical, cognitive, biological, and evolutionary perspectives. *Clinical Psychological Science*, 2167702616628523

30 Trimmer, P.C., *et al.* (2015) Adaptive learning can result in a failure to profit from good conditions: implications for understanding depression. *Evolution, medicine, and public health* 2015, 123-135
31 Gilbert, P. (2006) Evolution and depression: issues and implications. *Psychological medicine* 36, 287-297
32 Allen, N.B. and Badcock, P.B. (2003) The social risk hypothesis of depressed mood: evolutionary, psychosocial, and neurobiological perspectives. *Psychological bulletin* 129, 887
33 Vialou, V., *et al.* (2013) Epigenetic mechanisms of depression and antidepressants action. *Annual review of pharmacology and toxicology* 53, 59
34 Weissman, M.M., *et al.* (2005) Families at high and low risk for depression: a 3-generation study. *Archives of general psychiatry* 62, 29-36
35 Sun, H., *et al.* (2013) Epigenetics of the depressed brain: role of histone acetylation and methylation. *Neuropsychopharmacology* 38, 124-137
36 Meaney, M.J. (2001) Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annual review of neuroscience* 24, 1161-1192
37 Heim, C. and Binder, E.B. (2012) Current research trends in early life stress and depression: Review of human studies on sensitive periods, gene–environment interactions, and epigenetics. *Experimental neurology* 233, 102-111
38 Gold, P. (2015) The organization of the stress system and its dysregulation in depressive illness. *Molecular psychiatry* 20, 32-47
39 De Raedt, R. and Koster, E.H. (2010) Understanding vulnerability for depression from a cognitive neuroscience perspective: A reappraisal of attentional factors and a new conceptual framework. *Cognitive, Affective, & Behavioral Neuroscience* 10, 50-70
40 Slavich, G.M. and Irwin, M.R. (2014) From stress to inflammation and major depressive disorder: A social signal transduction theory of depression. *Psychological bulletin* 140, 774
41 Silk, J.S., *et al.* (2012) Why do anxious children become depressed teenagers? The role of social evaluative threat and reward processing. *Psychological medicine* 42, 2095-2107
42 Andersen, S.L. and Teicher, M.H. (2008) Stress, sensitive periods and maturational events in adolescent depression. *Trends in neurosciences* 31, 183-191
43 Gotlib, I.H. and Hammen, C. (2014) *Handbook of depression (3rd edn.)*. Guilford Press
44 Vrshek-Schallhorn, S., *et al.* (2014) Refining the candidate environment interpersonal stress, the serotonin transporter polymorphism, and gene-environment interactions in major depression. *Clinical Psychological Science* 2, 235-248
45 Cohen, S. and Wills, T.A. (1985) Stress, social support, and the buffering hypothesis. *Psychological bulletin* 98, 310
46 Hawkley, L.C. and Capitanio, J.P. (2015) Perceived social isolation, evolutionary fitness and health outcomes: a lifespan approach. *Phil. Trans. R. Soc. B* 370, 20140114
47 Leppänen, J.M. (2006) Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. *Current opinion in psychiatry* 19, 34-39

48 Hankin, B.L., *et al.* (2010) Corumination, interpersonal stress generation, and internalizing symptoms: Accumulating effects and transactional influences in a multiwave study of adolescents. *Development and psychopathology* 22, 217-235

49 Moore, M.T. and Fresco, D.M. (2012) Depressive realism: a meta-analytic review. *Clinical psychology review* 32, 496-509

50 Forgas, J.P. (2016) Can Sadness Be Good for You? *Australian Psychologist*

51 Pizzagalli, D.A. (2014) Depression, stress, and anhedonia: toward a synthesis and integrated model. *Annual review of clinical psychology* 10, 393

52 Hagen, E.H. (2011) Evolutionary theories of depression: a critical review. *The Canadian Journal of Psychiatry* 56, 716-726

53 Hames, J.L., *et al.* (2013) Interpersonal processes in depression. *Annual review of clinical psychology* 9, 355-377

54 Sloman, L. and Gilbert, P. (2000) *Subordination and defeat: An evolutionary approach to mood disorders and their therapy*. Routledge

55 Price, J.L. and Drevets, W.C. (2012) Neural circuits underlying the pathophysiology of mood disorders. *Trends in cognitive sciences* 16, 61-71

56 Kupferberg, A., *et al.* (2016) Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews* 69, 313-332

57 Rushworth, M.F., *et al.* (2013) Are there specialized circuits for social cognition and are they unique to humans? *Current opinion in neurobiology* 23, 436-442

58 Davey, C.G., *et al.* (2008) The emergence of depression in adolescence: Development of the prefrontal cortex and the representation of reward. *Neuroscience & Biobehavioral Reviews* 32, 1-19

59 Friston, K., *et al.* (2015) Active inference and epistemic value. *Cogn Neurosci*, 1-28

60 Friston, K., *et al.* (2014) The anatomy of choice: dopamine and decision-making. *Phil. Trans. R. Soc. B* 369, 20130481

61 Paulus, M.P. and Stein, M.B. (2006) An insular view of anxiety. *Biological psychiatry* 60, 383-387

62 Pezzulo, G., *et al.* (2015) Active Inference, homeostatic regulation and adaptive behavioural control. *Progress in neurobiology* 134, 17-35

63 Waselus, M., *et al.* (2011) Collateralized dorsal raphe nucleus projections: a mechanism for the integration of diverse functions during stress. *Journal of chemical neuroanatomy* 41, 266-280

64 Pellicano, E. and Burr, D. (2012) When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends in cognitive sciences* 16, 504-510

65 Fletcher, P.C. and Frith, C.D. (2009) Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci.* 10, 48-58

66 Friston, K.J., *et al.* (2014) Computational psychiatry: the brain as a phantastic organ. *The Lancet Psychiatry* 1, 148-158

67 Cuijpers, P., *et al.* (2011) Interpersonal psychotherapy for depression: a meta-analysis. *American Journal of Psychiatry* 168, 581-592

68 Park, H.-J. and Friston, K. (2013) Structural and functional brain networks: from connections to cognition. *Science* 342, 1238411

69 Hohwy, J. (2013) *The predictive mind*. Oxford University Press

70 Sporns, O. and Betzel, R.F. (2016) Modular brain networks. *Annual review of psychology* 67, 613-640

71 Buckner, R.L. and Krienen, F.M. (2013) The evolution of distributed association networks in the human brain. *Trends in Cognitive Sciences* 17, 648-665

72 Finlay, B.L. and Uchiyama, R. (2015) Developmental mechanisms channeling cortical evolution. *Trends in neurosciences* 38, 69-76

73 Anderson, M.L. and Finlay, B.L. (2014) Allocating structure to function: the strong links between neuroplasticity and natural selection. *Frontiers in human neuroscience* 7, 918

74 Mengistu, H.*, et al.* (2016) The evolutionary origins of hierarchy. *PLOS Comput Biol* 12, e1004829

75 Hilgetag, C.C. and Hütt, M.-T. (2014) Hierarchical modular brain connectivity is a stretch for criticality. *Trends in cognitive sciences* 18, 114-115

76 Hesse, J. and Gross, T. (2014) Self-organized criticality as a fundamental property of neural systems. *Frontiers in Systems Neuroscience* 8, 46-59

77 Sowell, E.R.*, et al.* (2003) Mapping cortical change across the human life span. *Nature neuroscience* 6, 309-315

78 Blakemore, S.-J. (2012) Imaging brain development: the adolescent brain. *Neuroimage* 61, 397-406

79 Spear, L.P. (2000) The adolescent brain and age-related behavioral manifestations. *Neuroscience & Biobehavioral Reviews* 24, 417-463

80 Fuhrmann, D.*, et al.* (2015) Adolescence as a sensitive period of brain development. *Trends in cognitive sciences* 19, 558-566

81 Blakemore, S.-J. and Mills, K.L. (2014) Is adolescence a sensitive period for sociocultural processing? *Annual review of psychology* 65, 187-207

82 Crone, E.A. and Dahl, R.E. (2012) Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature Reviews Neuroscience* 13, 636-650

83 Brown, B.B. and Larson, J. (2009) Peer relationships in adolescence. In *Handbook of Adolescent Psychology.* (Lerner, R.M. and Steinberg, L., eds), pp. 74-103, John Wiley & Sons

84 Sisk, C.L. and Zehr, J.L. (2005) Pubertal hormones organize the adolescent brain and behavior. *Frontiers in neuroendocrinology* 26, 163-174

86.