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)

1

1   **ABSTRACT**

2   Major depressive disorder is a debilitating condition characterised by diverse

3   neurocognitive and behavioural deficits. Nevertheless, our species-typical capacity for

4   depressed mood implies that it serves an adaptive function. Here, we apply an

5   interdisciplinary theory of brain function to explain depressed mood and its clinical

6   manifestations. Combining insights from the free-energy principle with evolutionary

7   theorising in psychology, we argue that depression reflects an adaptive response to

8   perceived threats of aversive social outcomes (e.g., exclusion) that minimises the

9   likelihood of surprising interpersonal exchanges (i.e., those with unpredictable

10   outcomes). We suggest that psychopathology typically arises from ineffectual

11   attempts to alleviate interpersonal difficulties and/or hyper-reactive neurobiological

12   responses to social stress (i.e., uncertainty), which often stems from early experience

13   that social uncertainty is difficult to resolve.

14

15   **Keywords:** Active Inference; Evolutionary Systems Theory; Depression; Free-

16   Energy Principle; Major Depressive Disorder

2

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

42   work with psychological research on the adaptive function of depression, along with

43   the familial, developmental and psychobiological mechanisms that often underlie it.

44   We propose that our species-typical capacity for depressed mood can be explained as

45   an evolved biobehavioural strategy that responds adaptively to adverse interpersonal

46   conditions by minimising the likelihood of unpredictable social interactions. We

47   discuss how our model builds on theories of clinical depression in the active inference

48   literature, before turning to the hierarchical neural mechanics that underlie depressed

49   mood and depressive disorder.

50

51   **Applying the Free-Energy Principle to Depression**

52   The FEP is a global theory of neural structure and function suggesting the brain can

53   be seen as a "prediction machine" that attempts to maximise the evidence for a

54   creature's model of the world by minimising an upper limit on surprise [i.e., free-

55   energy; see Box 2]. In line with predictive coding, the FEP describes the brain as a

56   hierarchical generative model – a hierarchy of hypotheses about the world that

57   enables a reduction of surprise by minimising discrepancies between incoming

58   sensory inputs and top-down predictions [9]. Conditional expectations are thought to

59   be encoded by deep pyramidal cells (i.e., representation units) at each level of the

60   cortical hierarchy that convey predictions downward to suppress errors at the level

61   below, while prediction errors are encoded by superficial pyramidal cells (i.e., error

62   units) that convey errors forward to revise expectations at the level above [10]. This

63   allows us to minimise surprise by updating our internal models (i.e., perception).

64   Alternatively, we can selectively sample sensory data to ensure that our predictions

65   are self-fulfilling – by changing how we act upon the world to confirm our

66   expectations (i.e., active inference [11]). Thus, perception and action operate

67   synergistically to minimise prediction errors and optimise our internal representations

68   of the environment. A key corollary of this model is that our predictions are optimised

69   by evolution, development and learning. Emphasis is placed on adaptive priors –

70   inherited expectations about the way our world unfolds that have been shaped by

71   natural selection to guide action-perception cycles toward adaptive (i.e., unsurprising)

72   states [5, 12].

73

74   To date, applications of the FEP to depressive disorders have chiefly concentrated on

75   two processes, stemming from different levels of the cortical hierarchy. The first

76   relates to limbic deficits in minimising prediction error. Barrett and colleagues

77   suggest that depressive disorders arise from aberrant interoceptive predictions

78   originating from abnormalities within the (limbic) agranular visceromotor cortex,

79   which is central to emotional processing, energy regulation and allostatic responses to

80   stress [13, 14]. These abnormalities can arise from past exposure to sustained distress,

81   and generate false (interoceptive) predictions about the body's upcoming autonomic,

82   metabolic and immunological needs that chronically activate physiological stress

83   responses (e.g., dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis and

84   pro-inflammatory states). Over time, the visceromotor systems try to minimise

85   prediction error by producing sickness behaviours (e.g., negative affect and fatigue;

86   also see [15]) that reduce energy expenditure and ultimately manifest in depression

87   [13]. Here, depression is seen as a disorder of allostasis, characterised by energy

88   dysregulation and deficits in interoceptive inference; i.e., an insensitivity to prediction

89   errors and/or a miscalibration of their precision – see glossary [14]. These deficits

90   lead to a failure to update dysfunctional internal models (e.g., cognitive rigidity),

91 perpetuating further metabolic inefficiencies and engendering the downward spiral

92 that typifies depressive illness.

93

94 The second class of models concentrates on impairments in top-down expectations of

95 reward. Checkroud [16] has described the depressed brain as a hierarchical

96 constellation of depressive beliefs, which impose a consistent negative bias in

97 predictions that manifests in anhedonic features and the down-regulation of neural

98 reward systems (e.g., dopaminergic and serotonergic dysfunction). In line with active

99 inference, these beliefs exacerbate depression by prompting the individual to actively

100 sample the environment to confirm negative predictions (e.g., learned helplessness).

101 Others have suggested that depressive disorders impair reward-approach behaviours

102 by causing a pathological underconfidence in one's predictions [17], or by distorting

103 higher-order evaluations of the self (e.g., low self-worth), disrupting social behaviour

104 by overweighting the likelihood of aversive interactions [18]. Each of these proposals

105 echo models of reinforcement learning in computational psychiatry and evolutionary

106 biology suggesting that depression emerges from successive discrepancies between

107 actual and expected reward outcomes (i.e., prediction errors), entrenching (empirical)

108 prior beliefs that rewards are unlikely which inhibit reward-approach behaviours [19,

109 20].

110

111 Taken together, the frameworks considered here suggest that depression entails

112 impairments in reward-approach systems emerging from two neurocognitive

113 processes – deficits in the predictive processing of sensory evidence; and prior beliefs

114 that negatively bias predictions. Although this perspective of depression as false

115 inference offers a cohesive, neurobiologically plausible account of the biobehavioural

116   deficits observed in depressive illness, two important questions remain. First, by

117   concentrating on depressive disorders, the models above say little about our species-

118   typical capacity for depressed or low mood. The notable exception is a formal

119   (computational) scheme that defines emotional valence in terms of the rate of change

120   in free-energy over time, with positive and negative affect tracking a decrease and

121   increase in free-energy, respectively (see [17]). In this model, negative moods enable

122   an organism to respond adaptively to unexpected changes in the world by increasing

123   the (learning) rate of evidence accumulation – overweighing recent sensory inputs

124   over prior experiences to heighten sensitivity to environmental change, thereby

125   minimising prediction error [17]. However, this does not specifically address the

126   adaptive significance of depression per se. Second, the literature to date sheds little

127   light on the ecological conditions responsible for the positive selective pressure that

128   depression appears to have [see Box 3]. If depression instantiates an adaptive prior, it

129   should minimise surprise in response to specific environmental challenges that have

130   threatened our inclusive fitness (i.e., free-energy) over evolutionary time. Identifying

131   this adaptive function is arguably central to understanding why depression occurs. To

132   address these issues, we hope to build upon the active inference literature by

133   incorporating complementary insights drawn from an evolutionary systems approach

134   to psychology.

135

136   **Insights from an Evolutionary Systems Approach to Psychology**

137   In psychology, evolutionary systems models have typically focused on the dynamic

138   interplay between evolutionary and developmental processes (e.g., [8, 21-25]), an

139   approach that has been further extended to reconcile theoretical divisions between

140   major paradigms in the field (see [4]). According to this perspective, the embodied

141 brain and behaviour emerge from selection acting on dynamic interactions between

142 the levels of causation identified by Tinbergen – adaptation, phylogeny, ontogeny,

143 and mechanism [26]. This causal hierarchy is arguably recapitulated by paradigms in

144 psychology, which concentrate differentially on four overlapping levels of

145 explanation – ultimate hypotheses for adaptive, species-typical characteristics (i.e.,

146 evolutionary psychology); epigenetic explanations for intergenerational, between-

147 group differences (i.e., evolutionary developmental biology and psychology);

148 dynamical explanations for individual similarities and differences (i.e., developmental

149 psychology); and mechanistic explanations for real-time phenomena (i.e.,

150 psychological subdisciplines such as cognitive, biological, personality, social and

151 clinical psychology) [4]. Central to EST is the need to explore how these causal levels

152 interact – evolutionary influences on neural structure and function constrain

153 individual development and learning, while effects at these more proximate levels can

154 shape the evolution of the brain [3, 27]. To explain depression then, we require a

155 multi-level hypothesis that synthesises diverse fields of psychological inquiry to

156 explain both why it is adaptive, along with how it emerges from intergenerational,

157 developmental and real-time mechanisms.

158

159 Although there are various Darwinian models of depression [28], a theme common to

160 many of these is that low mood reflects an adaptive biobehavioural strategy that

161 conserves or reallocates energy and resources in unpropitious social environments [29,

162 30]. According to this view, depressed mood states are elicited by aversive

163 interpersonal outcomes (e.g., exclusion, defeat, or loss) that indicate a critical loss of

164 control over social relationships that were critical to ancestral fitness [31]. A model

165 that incorporates influential theories in this area and shows promising conceptual

166   parallels with the FEP is the social risk hypothesis (technically, risk corresponds to

167   uncertainty and uncertainty is expected surprise or free energy). This maintains that

168   depressed mood reflects an adaptive, risk-averse approach to social interaction that

169   reduces the likelihood of further aversive outcomes by: (1) increasing our cognitive

170   sensitivity to (sensory) cues of social risk; (2) reducing our (behavioural) propensity

171   for taking social risks; and (3) initiating signalling behaviours that elicit support and

172   defuse conflict [32].

173

174   The idea that depression reflects an evolved response to adverse social conditions

175   concords with evidence that extends across Tinbergen's remaining levels of inquiry.

176   The intergenerational transmission of susceptibility to depressive disorders due to

177   deleterious social environments is widely documented [33, 34], with studies involving

178   rodents, primates, and humans showing that exposure to social stressors (e.g., low

179   maternal care and social defeat) produces potentially heritable epigenetic changes that

180   confer risk for disorder by heightening stress reactivity [35, 36]. Ontogenetically,

181   exposure to early social stress (e.g., parental loss, abuse, or neglect) is a strong

182   predictor of depressive vulnerability [37], and is thought to heighten susceptibility to

183   disorder by leading to hyperactivity of the HPA axis [38, 39] and up-regulating pro-

184   inflammatory immune responses [40]. Behavioural and neuroimaging studies further

185   suggest that the risk of depressive onset rises markedly in adolescence because of an

186   increased sensitivity to social threats in this period [41, 42]. Finally, research across

187   the sub-disciplines highlights an intimate connection between depression and the

188   social world (see [43]). For example, the precipitants of depression are typically

189   interpersonal in nature [44]; social support and belonging are key protective factors

190   [45]; and typical correlates of depression clearly exemplify negative self-other

191   relations (e.g., low self-esteem [46]). Consistent with the social risk hypothesis, there

192   are also multiple lines of evidence to suggest that low mood is associated with

193   biobehavioural changes that facilitate adaptive responses to social stress. Depressive

194   cognition is characterised by a specific, attentional bias towards socially-threatening

195   stimuli [47] and increased rumination about interpersonal problems [48], while

196   normative depressed states have been shown to increase the accuracy of social

197   inferences (e.g., depressive realism [49]) and improve social problem-solving [50].

198   Furthermore, many features of depression – such as anhedonia, a negative thinking

199   bias and social withdrawal – reduce exposure to social risks by inhibiting reward-

200   approach behaviours [51], while the signalling behaviours associated with depression

201   (e.g., reassurance-seeking and submissive behaviours) explicitly attempt to elicit

202   support and defuse potential conflict [52-54]. Notably, other studies have provided

203   direct empirical support for the social risk hypothesis itself (Badcock & Allen, 2003;

204   Badcock & Allen, 2007; Dunn, Whelton & Sharpe, 2012; Girard, Cohn, Mahoor,

205   Mavadati, & Rosenwald, 2013).

206

207   In light of such work, we suggest that the human capacity for depressed mood can be

208   explained in terms of a risk-averse adaptive prior that minimises uncertainty in the

209   social world when sensory cues indicate both a high degree of socio-environmental

210   volatility (i.e., unpredictability) and an increased probability of aversive interpersonal

211   outcomes (e.g., rejection, defeat or loss) (see Figure 1). This depressive response

212   instantiates a "better safe than sorry" strategy that minimises the likelihood of

213   unpredictable social interactions by causing adaptive changes in cognition (e.g.,

214   hypersensitivity to aversive social stimuli, a negative thinking bias and deficits in

215   responses to reward) and action (e.g., risk-averse behaviours such as social

216   withdrawal). Epigenetic and ontogenetic mechanisms arguably support this function

217   by sensitising the individual to socio-environmental volatility when developmental

218   insults indicate that the probability of aversive social interactions is high, producing

219   hyper-reactive stress response systems that heighten risk for disorder by increasing

220   the precision of social prediction errors and prompting exaggerated, pathological

221   responses to interpersonal stressors.

222

223   Notably, the exacerbation of normative depressed states into severe, dysfunctional

224   forms is also likely when depressive changes fail to alleviate social stress, creating a

225   self-perpetuating cycle arising from heightened and prolonged arousal of ineffectual

226   attempts to reduce socio-environmental volatility [32]. This, in turn, is likely to

227   chronically activate neurophysiological stress responses and leads to debilitating

228   sickness behaviours [13, 14]. As discussed above, previous applications of the FEP

229   suggest that this depressive spiral is engendered by a positive feedback loop between

230   two neurocognitive mechanisms – the increased precision of social prediction errors,

231   coupled with a negative bias affecting social predictions. Following active inference,

232   this is likely to engender ongoing depressive behaviours that seek to confirm negative

233   biases, creating a self-fulfilling prophecy (i.e., high predictability) born from mutually

234   reinforcing patterns of cognition and behaviour [16]. Here, depressed can be

235   interpreted as a maladaptive pattern of dysregulated defences – if this depressive

236   response is effective, an individual either escapes or avoids the social stressor or

237   adapts to it; if the defence fails, the individual is at risk of entering a self-perpetuating

238   dysregulated state, which falls beyond the normal range of adaptive functioning

239   (Allen & Badcock, 2003; Gilbert, 2001). Nevertheless, it should also be recognised

240   that clinical manifestations of depression can result from asocial causes that produce

241 neurobiological abnormalities typically associated with dysregulated mood (e.g., pro-

242 inflammatory immune responses induced by illness and medications; 40).

243

244 **The Depressed Brain**

245 It is widely accepted that depression emerges from bidirectional interactions between

246 hierarchically organised neural regions. Most of the theoretical work in this area

247 concentrates on two general brain systems that work in concert – a ventral affective

248 system, including subcortical regions such as the amygdala and ventral striatum; and

249 the prefrontal cortex (PFC), which modulates the reactions of the ventral affective

250 system [1]. These systems are composed of subcortical neural circuits responsible for

251 the unconscious processing of affective and social stimuli on the one hand; and on the

252 other, executive networks that regulate affective states, with medial prefrontal regions

253 playing a particularly important role in modulating visceral and behavioural responses

254 in order to adapt them to the external milieu [1].

255

256 More particularly, evidence gleaned from neuroimaging and animal studies suggests

257 that depression involves dysfunction of the "extended visceromotor network", in

258 which the medial PFC regulates affective states by modulating visceromotor output

259 via connections with the amygdala, ventral striatum, hypothalamus, and other

260 subcortical regions [55]. Brain regions across this network regulate motivation (e.g.,

261 anhedonia and dopaminergic function) and neurobiological responses to stress, and

262 play a central role in social threat and reward processing [39, 41, 56, 57].

263 Neurodevelopmental changes in these regions throughout adolescence are also

264 thought to heighten vulnerability to disorder by increasing sensitivity to rapidly

265 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.

290 Ultimately, our basic claim is that depression can be viewed as an adaptive faculty

291 that underwrites emotional allostasis in an increasingly prosocial and volatile world.

292 Physiologically, this faculty increases sensitivity to interpersonal, affiliative and

293 interoceptive cues. Clearly, sensitisation to stressful exteroceptive and interoceptive

294 cues also has to be predicted by the hierarchical brain, which implicates the functional

295 neuroanatomy described above. Under active inference, sensitivity to stress-related

296 cues corresponds to their precision [13, 60], implicating neuromodulatory systems

297 associated with reward, action selection and interoceptive inference [61-63].

298 Crucially, in order to act it is also necessary to attenuate the precision afforded to the

299 sensory consequences of action (i.e., we have to ignore the fact that we are not

300 currently acting). This means that an adaptive depressive response suspends sensory

301 attenuation – and action – so that we can attend to interpersonal prediction errors and

302 revise our (posterior) beliefs about our relationships with others, via perceptual

303 inference and learning. Sensory attenuation can be regarded as the complement of

304 sensory attention; i.e., attenuating or augmenting the gain (precision) afforded sensory

305 prediction errors to ignore or select sensory information, respectively. According to

306 this scheme, maladaptive forms of depression reflect a pervasive, self-maintaining

307 failure of sensory attenuation, leading to ruminations, false inference and a

308 concomitant inability to act and test these false beliefs.

309

310 Interestingly, exactly the same conclusions (namely, a failure of sensory attenuation)

311 have been drawn for a range of neuropsychiatric disorders, ranging from autism [64]

312 to schizophrenia [65]. One could ask what is specific about this mechanism in

313 depression, and respond by referring to the particular (interoceptive and affiliative)

314 modalities affected. However, perhaps the more intriguing implication is that the

315 comorbidity of depression and other disorders might arise from a common

316 pathophysiological mechanism, which can be explained in terms of false inference.

317

318 **Concluding Remarks**

319 In this opinion piece, we have endeavoured to contribute to the active inference

320 literature on mood disorder by suggesting that normative levels of depressed mood

321 instantiate an adaptive prior that minimises the likelihood of surprising interpersonal

322 interactions when faced with threats of aversive social outcomes that typically

323 compromised ancestral fitness. By extending beyond previous applications of the FEP

324 to emphasise both the adaptive function of low mood and the causal role of the social

325 ecology, we believe our model demonstrates the heuristic benefits of combining

326 active inference with insights in psychology to improve our understanding of

327 depressed mood and mood disorder. It also motivates new questions for research,

328 calling for greater integration between neuroscientific and psychological approaches

329 to explore the ways in which the neural mechanisms that underpin depression relate to

330 behaviour, development and the social world [see Outstanding Questions]. In

331 particular, the idea that depression can emerge from the need to navigate social risks

332 stands to inform theory-driven approaches in computational psychiatry, which

333 improve our understanding, prediction and treatment of mental illness by using

334 simulations and mathematical models to capture complex interactions across multiple

335 causal levels [19, 66].

336

337 That said, we do not wish to imply that depression is solely attributable to social

338 causes. In evolutionary psychology, for instance, the distinction between social and

339 non-social depressive responses is widely recognised [Durisko 2015; Gilbert 2006],

15

340 and as we have noted, depression can also arise from depressogenic neuroanatomical

341 abnormalities produced by influences other than unfavourable social conditions.

342 Nevertheless, our model adds to the active inference literature by emphasising the

343 importance of the social environment in explaining the aetiology and phenomenology

344 of depression. This underscores the need to develop (computational) diagnostic tools

345 that are capable of distinguishing between social and non-social forms in order to

346 inform treatment decisions.

347

348 In closing, our model also promotes clear avenues for intervention. To date,

349 proponents of the FEP have advocated treatments that directly target dysregulated

350 neural systems, such as psychopharmacological agents that act upon the

351 neurotransmitter systems that encode precision or uncertainty (e.g., serotonin and

352 dopamine [16]). They have also recommended the use of cognitive behavioural

353 therapies to disrupt the spiral of self-defeating actions typical of depression [16], or to

354 construct new prediction signals that modify the gain on prediction errors via the

355 salience network [14]. Our own model adds to this work by emphasising the need to

356 facilitate adaptive responses to social stress. This could well explain the efficacy of

357 interpersonal psychotherapy as a treatment for major depressive disorder [67], and

358 highlights the value of prevention and early intervention efforts that reduce

359 vulnerability by targeting modifiable risk factors in the social environment. Given the

360 heterogeneous nature of depression, we also recommend the development of

361 (computational) diagnostic tools capable of distinguishing between social and non-

362 social forms in order to inform treatment decisions. Finally, simply having a positive

363 and principled framework within which to understand depression – and the rationale

364 for therapeutic interventions – is likely to be helpful for those seeking treatment. Our

365 synthesis can be used to help clients understand why they have depression, and to

366 explain why, for example, it might be useful to combine interpersonal psychotherapy

367 with antidepressants.

368

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.

397 **Generative Model:** A probabilistic mapping from hidden causes in the environment

398 to observed consequences (sensory data), typically specified in terms of the likelihood

399 of observing some data (given their causes) and priors (on these causes).

400 **Interoception:** The perception and integration of autonomic, hormonal, visceral and

401 immunological (bodily) signals.

402 **Precision:** The inverse variance, volatility, or reliability of a signal. In predictive

403 coding, prediction errors are weighted by precisions that determine the relative

404 influence of bottom-up (error) and top-down (representation) signals (e.g., a high

405 precision on error signals corresponds to low confidence in top-down beliefs).

406 Dynamic precision weighting is mediated by neuromodulation and underwrites

407 psychological processes such as attentional selection and sensory attenuation.

408 **Predictive coding:** A processing scheme for inferring the likely causes of sensory

409 data by minimising prediction error. Typically, this entails a hierarchical generative

410 model (e.g., the brain) in which top-down signals convey predictions and bottom-up

411 signals convey (precision weighted) prediction errors.

412 **Prior:** The probability distribution or density on the causes of data that encode beliefs

413 about those causes prior to observing the data.

414 **Surprise**: The negative log probability of sensory experiences encountered by an

415 agent. Also known as surprisal or self-information.

416 **Visceromotor cortex**: Agranular (limbic) regions of isocortex and allocortex that

417 regulate the hormonal, immune and autonomic nervous systems, including the

418 cingulate cortex, the posterior ventral medial prefrontal cortex, the posterior

419 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].

444   It is now broadly accepted that a hierarchical neural structure is favoured by selection.

445   It enhances evolvability because deleterious changes to a single component of the

446   system are unlikely to affect the system itself, and it allows adaptive novelties to

447   emerge without disrupting global functioning [70]. Computer simulations of evolving

448   networks have also shown that selection favours a hierarchical organisation because it

449   conserves the (spatial, processing and metabolic) cost of neural connections; improves

450   problem-solving by recursively combining solutions to sub-problems; and adapts

451   more rapidly to new environments than non-hierarchical structures [74]. Finally, the

452   hierarchical brain is thought to promote "self-organised criticality". This is a

453   dynamical state poised between completely ordered, stable cycles of activity and

454   highly complex, chaotic ones that optimises evolvability because it allows small

455   extrinsic changes to elicit large intrinsic reorganisations. The hierarchical segregation

456   of neural networks into distributed neighbourhoods has been found to stretch the

457   parameter range for self-organised criticality by allowing subcritical and supercritical

458   dynamics to co-exist simultaneously [75]. Since systems at criticality have optimal

459   information-processing capacities, a structure that extends this critical region is likely

460   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,

487   action and perception operate synergistically to optimise an organism's model of the

488   environment. Crucially, to minimise free-energy, the precision of prediction errors

489   also has to be predicted, invoking notions of attentional gain (psychologically) and

490   neuromodulation (physiologically).

491

492   The FEP also applies to the morphology, development and evolution of the brain. It

493   suggests that instead of just containing a model of the world, the brain is a model of

494   the world – a physical transcription of causal regularities in the environment that is

495   optimised by evolution. This model instantiates genetically specified (empirical) prior

496   beliefs that have minimised free-energy (i.e., maximised model evidence) over

497   evolutionary time by ensuring an organism seeks out a small number of unsurprising

498   states that are consistent with its phenotype and environment. In other words, natural

499   selection is nature's way of performing Bayesian model selection to minimise the

500   (variational) free energy of phenotypes (i.e., generative models).

501

502   **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

528   social cues. Moreover, prefrontal cortical development leads, on the one hand, to

529   improved regulation of affective processes, but on the other, heightens sensitivity to

530   the nuance and complexity of interpersonal relationships [58]. For this reason,

531   increased vulnerability to depression starts in puberty but is maintained well beyond

532   adolescence.

533

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.

559   13. Barrett, L.F. and Simmons, W.K. (2015) Interoceptive predictions in the brain.

560   Nat. Rev. Neurosci. 16, 419-429.

561   14. Barrett, L.F. et al. (2016) An active inference theory of allostasis and

562   interoception in depression. Philos. Trans. R. Soc. Lond. B Biol. Sci. 371,

563   20160011.

564   15. Seth, A.K. and Friston, K.J. (2016) Active interoceptive inference and the

565   emotional brain. Philos. Trans. R. Soc. Lond. B Biol. Sci. 371, 20160007.

566   16. Chekroud, A.M. (2015) Unifying treatments for depression: An application of the

567   Free Energy Principle. Front Psychol 6, 153.

568   17. Joffily, M. and Coricelli, G. (2013) Emotional valence and the free-energy

569   principle. PLoS Comput. Biol. 9, e1003094.

570   18. Moutoussis, M. et al. (2014) Bayesian inferences about the self (and others): A

571   review. Conscious. Cogn. 25, 67-76.

572   19. Adams, R.A. et al. (2016) Computational psychiatry: Towards a mathematically

573   informed understanding of mental illness. J. Neurol. Neurosurg. Psychiatry 87,

574   53-63.

575   20. Nettle, D. and Bateson, M. (2012) The evolutionary origins of mood and its

576   disorders. Current Biology 22, R712-R721.

577   21. Caporael, L.R. (2001) Evolutionary psychology: Toward a unifying theory and a

578   hybrid science. Annu. Rev. Psychol. 52, 607-628.

579   22. Geary, D.C. and Bjorklund, D.F. (2000) Evolutionary developmental psychology.

580   Child Dev. 71, 57-65.

581   23. Kenrick, D.T. et al. (2002) Dynamical evolutionary psychology: Mapping the

582   domains of the new interactionist paradigm. Pers. Soc. Psychol. Rev. 6, 347-356.

583   24. Ploeger, A. et al. (2008) Is evolutionary psychology a metatheory for psychology?

584   A discussion of four major issues in psychology from an evolutionary

585   developmental perspective. Psychol. Inq. 19, 1-18.

586   25. Frankenhuis, W.E. et al. (2013) Bridging developmental systems theory and

587   evolutionary psychology using dynamic optimization. Dev. Sci. 16, 584-598.

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

589   433.

590   27. Marshall, P.J. (2013) Coping with complexity: Developmental systems and

591   multilevel analyses in developmental psychopathology. Dev. Psychopathol. 25,

592   1311-1324.

593   28. Durisko, Z. et al. (2015) An adaptationist perspective on the etiology of

594   depression. J. Affect. Disord. 172, 315-323.

595   29. Beck, A.T. and Bredemeier, K. (2016) A unified model of depression integrating

596   clinical, cognitive, biological, and evolutionary perspectives. Clin. Psychol. Sci. 4,

597   596-619.

598   30. Trimmer, P.C. et al. (2015) Adaptive learning can result in a failure to profit from

599   good conditions: Implications for understanding depression. Evol. Med. Public

600   Health 2015, 123-135.

601   31. Gilbert, P. (2006) Evolution and depression: Issues and implications. Psychol.

602   Med. 36, 287-297.

603   32. Allen, N.B. and Badcock, P.B. (2003) The social risk hypothesis of depressed

604   mood: Evolutionary, psychosocial, and neurobiological perspectives. Psychol.

605   Bull. 129, 887-913.

606   33. Vialou, V. et al. (2013) Epigenetic mechanisms of depression and antidepressants

607   action. Annu. Rev. Pharmacol. Toxicol. 53, 59-87.

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

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

612   36. Meaney, M.J. (2001) Maternal care, gene expression, and the transmission of
613   individual differences in stress reactivity across generations. Annu. Rev. Neurosci.
614   24, 1161-1192.

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

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

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

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

626   41. Silk, J.S. et al. (2012) Why do anxious children become depressed teenagers? The
627   role of social evaluative threat and reward processing. Psychol. Med. 42, 2095-
628   2107.

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

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

633   44. Vrshek-Schallhorn, S. et al. (2014) Refining the candidate environment

634   interpersonal stress, the serotonin transporter polymorphism, and gene-

635   environment interactions in major depression. Clin. Psychol. Sci. 2, 235-248.

636   45. Cohen, S. and Wills, T.A. (1985) Stress, social support, and the buffering

637   hypothesis. Psychol. Bull. 98, 310-357.

638   46. Hawkley, L.C. and Capitanio, J.P. (2015) Perceived social isolation, evolutionary

639   fitness and health outcomes: A lifespan approach. Philos. Trans. R. Soc. Lond. B

640   Biol. Sci. 370, 20140114.

641   47. Leppänen, J.M. (2006) Emotional information processing in mood disorders: A

642   review of behavioral and neuroimaging findings. Curr. Opin. Psychiatry 19, 34-39.

643   48. Hankin, B.L. et al. (2010) Corumination, interpersonal stress generation, and

644   internalizing symptoms: Accumulating effects and transactional influences in a

645   multiwave study of adolescents. Dev. Psychopathol. 22, 217-235.

646   49. Moore, M.T. and Fresco, D.M. (2012) Depressive realism: A meta-analytic

647   review. Clin. Psychol. Rev. 32, 496-509.

648   50. Forgas, J.P. (2016) Can sadness be good for you? Aust. Psychol.

649   DOI: 10.1111/ap.12230

650   51. Pizzagalli, D.A. (2014) Depression, stress, and anhedonia: Toward a synthesis and

651   integrated model. Annu. Rev. Clin. Psychol. 10, 393.

652   52. Hagen, E.H. (2011) Evolutionary theories of depression: A critical review. Can. J.

653   Psychiatry 56, 716-726.

654   53. Hames, J.L. et al. (2013) Interpersonal processes in depression. Annu. Rev. Clin.

655   Psychol. 9, 355-377.

656   54. Sloman, L. and Gilbert, P. (2000) *Subordination and Defeat: An Evolutionary*

657   *approach to Mood Disorders and Their Therapy*, Routledge.

658   55. Price, J.L. and Drevets, W.C. (2012) Neural circuits underlying the

659   pathophysiology of mood disorders. Trends Cogn. Sci. 16, 61-71.

660   56. Kupferberg, A. et al. (2016) Social functioning in major depressive disorder.

661   Neurosci. Biobehav. Rev. 69, 313-332.

662   57. Rushworth, M.F. et al. (2013) Are there specialized circuits for social cognition

663   and are they unique to humans? Curr. Opin. Neurobiol. 23, 436-442.

664   58. Davey, C.G. et al. (2008) The emergence of depression in adolescence:

665   Development of the prefrontal cortex and the representation of reward. Neurosci.

666   Biobehav. Rev. 32, 1-19.

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

668   28.

669   60. Friston, K. et al. (2014) The anatomy of choice: Dopamine and decision-making.

670   Philos. Trans. R. Soc. Lond. B Biol. Sci. 369, 20130481.

671   61. Paulus, M.P. and Stein, M.B. (2006) An insular view of anxiety. Biol. Psychiatry

672   60, 383-387.

673   62. Pezzulo, G. et al. (2015) Active Inference, homeostatic regulation and adaptive

674   behavioural control. Prog. Neurobiol. 134, 17-35.

675   63. Waselus, M. et al. (2011) Collateralized dorsal raphe nucleus projections: A

676   mechanism for the integration of diverse functions during stress. J. Chem.

677   Neuroanat. 41, 266-280.

678   64. Pellicano, E. and Burr, D. (2012) When the world becomes 'too real': A Bayesian

679   explanation of autistic perception. Trends Cogn. Sci. 16, 504-510.

680   65. Fletcher, P.C. and Frith, C.D. (2009) Perceiving is believing: A Bayesian

681   approach to explaining the positive symptoms of schizophrenia. Nat. Rev.

682   Neurosci. 10, 48-58.

683   66. Friston, K.J. et al. (2014) Computational psychiatry: The brain as a phantastic

684   organ. Lancet Psychiatry 1, 148-158.

685   67. Cuijpers, P. et al. (2011) Interpersonal psychotherapy for depression: A meta-

686   analysis. Am. J. Psychiatry 168, 581-592.

687   68. Park, H.-J. and Friston, K. (2013) Structural and functional brain networks: from

688   connections to cognition. Sci. 342, 1238411.

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

690   70. Sporns, O. and Betzel, R.F. (2016) Modular brain networks. Annu. Rev. Psychol.

691   67, 613-640.

692   71. Buckner, R.L. and Krienen, F.M. (2013) The evolution of distributed association

693   networks in the human brain. Trends Cogn. Sci. 17, 648-665.

694   72. Finlay, B.L. and Uchiyama, R. (2015) Developmental mechanisms channeling

695   cortical evolution. Trends Neurosci. 38, 69-76.

696   73. Anderson, M.L. and Finlay, B.L. (2014) Allocating structure to function: The

697   strong links between neuroplasticity and natural selection. Front. Hum. Neurosci.

698   7, 918.

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

700   Biol. 12, e1004829.

701   75. Hilgetag, C.C. and Hütt, M.-T. (2014) Hierarchical modular brain connectivity is

702   a stretch for criticality. Trends Cogn. Sci. 18, 114-115.

703   76. Hesse, J. and Gross, T. (2014) Self-organized criticality as a fundamental property

704   of neural systems. Front. Syst. Neurosci. 8, 46-59.

705   77. Sowell, E.R. et al. (2003) Mapping cortical change across the human life span.

706   Nat. Neurosci. 6, 309-315.

707   78. Blakemore, S.-J. (2012) Imaging brain development: the adolescent brain.

708   Neuroimage 61, 397-406.

709   79. Spear, L.P. (2000) The adolescent brain and age-related behavioral manifestations.

710   Neurosci. Biobehav. Rev. 24, 417-463.

711   80. Fuhrmann, D. et al. (2015) Adolescence as a sensitive period of brain

712   development. Trends Cogn. Sci. 19, 558-566.

713   81. Blakemore, S.-J. and Mills, K.L. (2014) Is adolescence a sensitive period for

714   sociocultural processing? Annu. Rev. Psychol. 65, 187-207.

715   82. Crone, E.A. and Dahl, R.E. (2012) Understanding adolescence as a period of

716   social–affective engagement and goal flexibility. Nat. Rev. Neurosci. 13, 636-650.

717   83. Brown, B.B. and Larson, J. (2009) Peer relationships in adolescence. In Handbook

718   of Adolescent Psychology (3rd edn) (Lerner, R.M. and Steinberg, L. eds), pp. 74-

719   103, John Wiley & Sons.

720   84. Sisk, C.L. and Zehr, J.L. (2005) Pubertal hormones organize the adolescent brain

721   and behavior. Front. Neuroendocrinol. 26, 163-174.

722   85.

723   1 Pfeifer, J.H. and Allen, N.B. (2012) Arrested development? Reconsidering dual-
724   systems models of brain function in adolescence and disorders. *Trends in*
725   *cognitive sciences* 16, 322-329
726   2 Pfeifer, J.H. and Allen, N.B. (2016) The audacity of specificity: Moving
727   adolescent developmental neuroscience towards more powerful scientific
728   paradigms and translatable models. *Developmental cognitive neuroscience* 17,
729   131-137
730   3 Badcock, P.*, et al.* (submitted) The Hierarchically Mechanistic Mind: An
731   evolutionary systems theory of the brain and behaviour. .
732   4 Badcock, P.B. (2012) Evolutionary systems theory: A unifying meta-theory of
733   psychological science. *Review of General Psychology* 16, 10
734   5 Friston, K. (2010) The free-energy principle: a unified brain theory? *Nature*
735   *Reviews Neuroscience* 11, 127-138
736   6 Anderson, M.L. (2014) *After phrenology*. MIT Press
737   7 Clark, A. (2015) *Surfing uncertainty: Prediction, action, and the embodied mind*.
738   Oxford University Press
739   8 Lickliter, R. and Honeycutt, H. (2003) Developmental dynamics: toward a
740   biologically plausible evolutionary psychology. *Psychological bulletin* 129, 819

741 9 Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the
742 future of cognitive science. *Behavioral and Brain Sciences* 36, 181-204
743 10 Bastos, A.M.*, et al.* (2012) Canonical microcircuits for predictive coding.
744 *Neuron* 76, 695-711
745 11 Friston, K.J.*, et al.* (2010) Action and behavior: a free-energy formulation.
746 *Biological cybernetics* 102, 227-260
747 12 Friston, K. (2013) Life as we know it. *Journal of the Royal Society Interface* 10,
748 20130475
749 13 Barrett, L.F. and Simmons, W.K. (2015) Interoceptive predictions in the brain.
750 *Nature Reviews Neuroscience* 16, 419-429
751 14 Barrett, L.F.*, et al.* (2016) An active inference theory of allostasis and
752 interoception in depression. *Phil. Trans. R. Soc. B* 371, 20160011
753 15 Seth, A.K. and Friston, K.J. (2016) Active interoceptive inference and the
754 emotional brain. *Phil. Trans. R. Soc. B* 371, 20160007
755 16 Chekroud, A.M. (2015) Unifying treatments for depression: an application of
756 the Free Energy Principle. *Frontiers in psychology* 6
757 17 Joffily, M. and Coricelli, G. (2013) Emotional valence and the free-energy
758 principle. *PLoS Comput Biol* 9, e1003094
759 18 Moutoussis, M.*, et al.* (2014) Bayesian inferences about the self (and others):
760 A review. *Consciousness and cognition* 25, 67-76
761 19 Adams, R.A.*, et al.* (2016) Computational Psychiatry: towards a
762 mathematically informed understanding of mental illness. *Journal of Neurology,*
763 *Neurosurgery & Psychiatry* 87, 53-63
764 20 Nettle, D. and Bateson, M. (2012) The evolutionary origins of mood and its
765 disorders. *Current Biology* 22, R712-R721
766 21 Caporael, L.R. (2001) Evolutionary psychology: Toward a unifying theory and
767 a hybrid science. *Annual review of psychology* 52, 607-628
768 22 Geary, D.C. and Bjorklund, D.F. (2000) Evolutionary developmental
769 psychology. *Child development* 71, 57-65
770 23 Kenrick, D.T.*, et al.* (2002) Dynamical evolutionary psychology: Mapping the
771 domains of the new interactionist paradigm. *Personality and Social Psychology*
772 *Review* 6, 347-356
773 24 Ploeger, A.*, et al.* (2008) Is evolutionary psychology a metatheory for
774 psychology? A discussion of four major issues in psychology from an
775 evolutionary developmental perspective. *Psychological Inquiry* 19, 1-18
776 25 Frankenhuis, W.E.*, et al.* (2013) Bridging developmental systems theory and
777 evolutionary psychology using dynamic optimization. *Developmental Science* 16,
778 584-598
779 26 Tinbergen, N. (1963) On aims and methods of ethology. *Zeitschrift für*
780 *Tierpsychologie* 20, 410-433
781 27 Marshall, P.J. (2013) Coping with complexity: Developmental systems and
782 multilevel analyses in developmental psychopathology. *Development and*
783 *psychopathology* 25, 1311-1324
784 28 Durisko, Z.*, et al.* (2015) An adaptationist perspective on the etiology of
785 depression. *Journal of affective disorders* 172, 315-323
786 29 Beck, A.T. and Bredemeier, K. (2016) A unified model of depression
787 integrating clinical, cognitive, biological, and evolutionary perspectives. *Clinical*
788 *Psychological Science*, 2167702616628523

789   30 Trimmer, P.C.*, et al.* (2015) Adaptive learning can result in a failure to profit
790   from good conditions: implications for understanding depression. *Evolution,*
791   *medicine, and public health* 2015, 123-135
792   31 Gilbert, P. (2006) Evolution and depression: issues and implications.
793   *Psychological medicine* 36, 287-297
794   32 Allen, N.B. and Badcock, P.B. (2003) The social risk hypothesis of depressed
795   mood: evolutionary, psychosocial, and neurobiological perspectives.
796   *Psychological bulletin* 129, 887
797   33 Vialou, V.*, et al.* (2013) Epigenetic mechanisms of depression and
798   antidepressants action. *Annual review of pharmacology and toxicology* 53, 59
799   34 Weissman, M.M.*, et al.* (2005) Families at high and low risk for depression: a
800   3-generation study. *Archives of general psychiatry* 62, 29-36
801   35 Sun, H.*, et al.* (2013) Epigenetics of the depressed brain: role of histone
802   acetylation and methylation. *Neuropsychopharmacology* 38, 124-137
803   36 Meaney, M.J. (2001) Maternal care, gene expression, and the transmission of
804   individual differences in stress reactivity across generations. *Annual review of*
805   *neuroscience* 24, 1161-1192
806   37 Heim, C. and Binder, E.B. (2012) Current research trends in early life stress
807   and depression: Review of human studies on sensitive periods, gene–
808   environment interactions, and epigenetics. *Experimental neurology* 233, 102-111
809   38 Gold, P. (2015) The organization of the stress system and its dysregulation in
810   depressive illness. *Molecular psychiatry* 20, 32-47
811   39 De Raedt, R. and Koster, E.H. (2010) Understanding vulnerability for
812   depression from a cognitive neuroscience perspective: A reappraisal of
813   attentional factors and a new conceptual framework. *Cognitive, Affective, &*
814   *Behavioral Neuroscience* 10, 50-70
815   40 Slavich, G.M. and Irwin, M.R. (2014) From stress to inflammation and major
816   depressive disorder: A social signal transduction theory of depression.
817   *Psychological bulletin* 140, 774
818   41 Silk, J.S.*, et al.* (2012) Why do anxious children become depressed teenagers?
819   The role of social evaluative threat and reward processing. *Psychological*
820   *medicine* 42, 2095-2107
821   42 Andersen, S.L. and Teicher, M.H. (2008) Stress, sensitive periods and
822   maturational events in adolescent depression. *Trends in neurosciences* 31, 183-
823   191
824   43 Gotlib, I.H. and Hammen, C. (2014) *Handbook of depression (3rd edn.).* Guilford
825   Press
826   44 Vrshek-Schallhorn, S.*, et al.* (2014) Refining the candidate environment
827   interpersonal stress, the serotonin transporter polymorphism, and gene-
828   environment interactions in major depression. *Clinical Psychological Science* 2,
829   235-248
830   45 Cohen, S. and Wills, T.A. (1985) Stress, social support, and the buffering
831   hypothesis. *Psychological bulletin* 98, 310
832   46 Hawkley, L.C. and Capitanio, J.P. (2015) Perceived social isolation,
833   evolutionary fitness and health outcomes: a lifespan approach. *Phil. Trans. R. Soc.*
834   *B* 370, 20140114
835   47 Leppänen, J.M. (2006) Emotional information processing in mood disorders: a
836   review of behavioral and neuroimaging findings. *Current opinion in psychiatry* 19,
837   34-39

838 48 Hankin, B.L., *et al.* (2010) Corumination, interpersonal stress generation, and
839 internalizing symptoms: Accumulating effects and transactional influences in a
840 multiwave study of adolescents. *Development and psychopathology* 22, 217-235
841 49 Moore, M.T. and Fresco, D.M. (2012) Depressive realism: a meta-analytic
842 review. *Clinical psychology review* 32, 496-509
843 50 Forgas, J.P. (2016) Can Sadness Be Good for You? *Australian Psychologist*
844 51 Pizzagalli, D.A. (2014) Depression, stress, and anhedonia: toward a synthesis
845 and integrated model. *Annual review of clinical psychology* 10, 393
846 52 Hagen, E.H. (2011) Evolutionary theories of depression: a critical review. *The*
847 *Canadian Journal of Psychiatry* 56, 716-726
848 53 Hames, J.L., *et al.* (2013) Interpersonal processes in depression. *Annual review*
849 *of clinical psychology* 9, 355-377
850 54 Sloman, L. and Gilbert, P. (2000) *Subordination and defeat: An evolutionary*
851 *approach to mood disorders and their therapy*. Routledge
852 55 Price, J.L. and Drevets, W.C. (2012) Neural circuits underlying the
853 pathophysiology of mood disorders. *Trends in cognitive sciences* 16, 61-71
854 56 Kupferberg, A., *et al.* (2016) Social functioning in major depressive disorder.
855 *Neuroscience & Biobehavioral Reviews* 69, 313-332
856 57 Rushworth, M.F., *et al.* (2013) Are there specialized circuits for social
857 cognition and are they unique to humans? *Current opinion in neurobiology* 23,
858 436-442
859 58 Davey, C.G., *et al.* (2008) The emergence of depression in adolescence:
860 Development of the prefrontal cortex and the representation of reward.
861 *Neuroscience & Biobehavioral Reviews* 32, 1-19
862 59 Friston, K., *et al.* (2015) Active inference and epistemic value. *Cogn Neurosci*,
863 1-28
864 60 Friston, K., *et al.* (2014) The anatomy of choice: dopamine and decision-
865 making. *Phil. Trans. R. Soc. B* 369, 20130481
866 61 Paulus, M.P. and Stein, M.B. (2006) An insular view of anxiety. *Biological*
867 *psychiatry* 60, 383-387
868 62 Pezzulo, G., *et al.* (2015) Active Inference, homeostatic regulation and
869 adaptive behavioural control. *Progress in neurobiology* 134, 17-35
870 63 Waselus, M., *et al.* (2011) Collateralized dorsal raphe nucleus projections: a
871 mechanism for the integration of diverse functions during stress. *Journal of*
872 *chemical neuroanatomy* 41, 266-280
873 64 Pellicano, E. and Burr, D. (2012) When the world becomes 'too real': a
874 Bayesian explanation of autistic perception. *Trends in cognitive sciences* 16, 504-
875 510
876 65 Fletcher, P.C. and Frith, C.D. (2009) Perceiving is believing: a Bayesian
877 approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci.*
878 10, 48-58
879 66 Friston, K.J., *et al.* (2014) Computational psychiatry: the brain as a phantastic
880 organ. *The Lancet Psychiatry* 1, 148-158
881 67 Cuijpers, P., *et al.* (2011) Interpersonal psychotherapy for depression: a meta-
882 analysis. *American Journal of Psychiatry* 168, 581-592
883 68 Park, H.-J. and Friston, K. (2013) Structural and functional brain networks:
884 from connections to cognition. *Science* 342, 1238411
885 69 Hohwy, J. (2013) *The predictive mind*. Oxford University Press

886 70 Sporns, O. and Betzel, R.F. (2016) Modular brain networks. *Annual review of*
887 *psychology* 67, 613-640
888 71 Buckner, R.L. and Krienen, F.M. (2013) The evolution of distributed
889 association networks in the human brain. *Trends in Cognitive Sciences* 17, 648-
890 665
891 72 Finlay, B.L. and Uchiyama, R. (2015) Developmental mechanisms channeling
892 cortical evolution. *Trends in neurosciences* 38, 69-76
893 73 Anderson, M.L. and Finlay, B.L. (2014) Allocating structure to function: the
894 strong links between neuroplasticity and natural selection. *Frontiers in human*
895 *neuroscience* 7, 918
896 74 Mengistu, H.*, et al.* (2016) The evolutionary origins of hierarchy. *PLOS Comput*
897 *Biol* 12, e1004829
898 75 Hilgetag, C.C. and Hütt, M.-T. (2014) Hierarchical modular brain connectivity
899 is a stretch for criticality. *Trends in cognitive sciences* 18, 114-115
900 76 Hesse, J. and Gross, T. (2014) Self-organized criticality as a fundamental
901 property of neural systems. *Frontiers in Systems Neuroscience* 8, 46-59
902 77 Sowell, E.R.*, et al.* (2003) Mapping cortical change across the human life span.
903 *Nature neuroscience* 6, 309-315
904 78 Blakemore, S.-J. (2012) Imaging brain development: the adolescent brain.
905 *Neuroimage* 61, 397-406
906 79 Spear, L.P. (2000) The adolescent brain and age-related behavioral
907 manifestations. *Neuroscience & Biobehavioral Reviews* 24, 417-463
908 80 Fuhrmann, D.*, et al.* (2015) Adolescence as a sensitive period of brain
909 development. *Trends in cognitive sciences* 19, 558-566
910 81 Blakemore, S.-J. and Mills, K.L. (2014) Is adolescence a sensitive period for
911 sociocultural processing? *Annual review of psychology* 65, 187-207
912 82 Crone, E.A. and Dahl, R.E. (2012) Understanding adolescence as a period of
913 social–affective engagement and goal flexibility. *Nature Reviews Neuroscience* 13,
914 636-650
915 83 Brown, B.B. and Larson, J. (2009) Peer relationships in adolescence. In
916 *Handbook of Adolescent Psychology.* (Lerner, R.M. and Steinberg, L., eds), pp. 74-
917 103, John Wiley & Sons
918 84 Sisk, C.L. and Zehr, J.L. (2005) Pubertal hormones organize the adolescent
919 brain and behavior. *Frontiers in neuroendocrinology* 26, 163-174
920 86.