# 1    In the Body's Eye

## 2    The Computational Anatomy of Interoceptive Inference

3          Micah Allen[1,2,3], Andrew Levy[4], Thomas Parr[4], Karl J. Friston[4]

4

5          [1]Aarhus Institute of Advanced Studies, Aarhus University, Denmark

6      [2]Centre of Functionally Integrative Neuroscience, Aarhus University Hospital, Denmark

7          [3]Cambridge Psychiatry, Cambridge University, Cambridge UK

8          [4]Wellcome Centre for Human Neuroimaging, UCL

9

**10    Abstract**

11

12    A growing body of evidence highlights the intricate linkage of exteroceptive perception to
13    the rhythmic activity of the visceral body. In parallel, interoceptive inference theories of
14    emotion and self-consciousness are on the rise in cognitive science. However, thus far no
15    formal theory has emerged to integrate these twin domains; instead most extant work is
16    conceptual in nature. Here, we introduce a formal model of cardiac active inference, which
17    explains how ascending cardiac signals entrain exteroceptive sensory perception and
18    confidence. Through simulated psychophysics, we reproduce the defensive startle reflex and
19    commonly reported effects linking the cardiac cycle to fear perception. We further show that
20    simulated 'interoceptive lesions' blunt fear expectations, induce psychosomatic
21    hallucinations, and exacerbate metacognitive biases. Through synthetic heart-rate
22    variability analyses, we illustrate how the balance of arousal-priors and visceral prediction
23    errors produces idiosyncratic patterns of physiological reactivity. Our model thus offers the
24    possibility to computationally phenotype disordered brain-body interaction.

25

26

**27    Introduction**

28       The enactive view of perception – implied by active vision and inference – suggests
29    an intimate co-dependency between perception and the active sampling of our sensorium.

30  In this work, we take the embodied view to its ultimate conclusion and consider perception
31  as a function of the physical and physiological body we use to 'measure' the world. In
32  particular, our focus is on the coupling – or interaction – between interoceptive and
33  exteroceptive perception; namely, how bodily states and states of affairs beyond the body
34  are inferred – and how inference about each domain affects the other. For example, does
35  what we see depend upon our autonomic status and how does visual perceptual synthesis
36  affect sympathetic or parasympathetic outflow? The body is, in essence, an ensemble of
37  fluctuating systems with biorhythms nested at multiple timescales. How then do these
38  physiological fluctuations interact with perceptual synthesis in the visual and auditory
39  domains?

40      There is a rapidly growing body of evidence suggesting that bodily and autonomic
41  states affect perceptual and metacognitive decisions (Allen et al., 2016b; Azevedo et al., 2017;
42  Bonvallet and Bloch, 1961; Cohen et al., 1980; Garfinkel et al., 2014; Hauser et al., 2017b;
43  Lacey and Lacey, 1978; Park et al., 2014; Salomon et al., 2016; Velden and Juris, 1975; Zelano
44  et al., 2016). Much of this evidence emphasises the dynamic aspect of our physiology; usually
45  assessed in terms of how psychophysics depends upon the phase of some physiological cycle.
46  Most of the empirical evidence suggests that biorhythms gate or modulate the way that
47  sensory evidence is accumulated during perception (Bonvallet et al., 1954; Bonvallet and
48  Bloch, 1961; Karavaev et al., 2018; Varga and Heck, 2017). In the predictive coding literature,
49  this is usually treated as fluctuating, context sensitive, changes in the precision of sensory
50  sampling (e.g., the precision or gain of prediction errors). Clear examples of this include the
51  fast waxing and waning of precision during active visual sampling. For example, saccadic
52  suppression – during saccadic eye movements – alternates with attention to fixated visual
53  information every 250 ms or so. This process of actively sampling the environment via
54  ballistic saccade itself varies with the cardiac cycle (Galvez-Pol et al., 2018; Kunzendorf et al.,
55  2019; Ohl et al., 2016). At still slower timescales, respiratory (Herrero et al., 2017; Tort et al.,
56  2018b, 2018a; Zelano et al., 2016) are coupled to neuronal oscillations and behavior. In
57  short, at probably every timescale there are systematic fluctuations in the precision or
58  quality of sensory evidence that depend upon when we actually interrogate the world, in
59  relation to the biorhythms of our sensory apparatus; namely, our body.

60    Our focus on the multimodal integration of interoceptive and exteroceptive domains
61    is driven by the overwhelming evidence for interoception as a key modality in hedonics,
62    arousal, emotion and selfhood (Allen and Friston, 2018; Apps and Tsakiris, 2014; Gallagher
63    and Allen, 2018; Seth, 2013; Seth and Friston, 2016). This is generally treated under the
64    rubric of interoceptive inference; namely, active inference in the interoceptive domain.
65    There are several compelling formulations of interoceptive inference from the perspective
66    of neurophysiology, neuroanatomy and, indeed, issues of consciousness in terms of minimal
67    selfhood. However, much of this treatment rests upon a purely conceptual analysis –
68    underpinned by some notion of active (Bayesian) inference about states of the world
69    (including the body). In this work, we offer a more formal (mathematical) analysis that we
70    hope will be a point of reference for both theoretical and empirical investigations.

71    In brief, we constructed a (minimal) active inference architecture to simulate
72    embodied perception and concomitant arousal. Here, we focused on simulating interactions
73    between the cardiac cycle and exteroceptive perception. In principle however, our
74    simulation provides a computational proof-of-principle that can be expanded to understand
75    brain-body coupling at any physiological or behavioral timescale. Using a Markov decision
76    process formulation, we created a synthetic subject who exhibited physiological (cardio-
77    acceleration) responses to arousing stimuli. Our agenda was twofold: first, to provide a
78    sufficiency proof that – in at least one example – the interaction between interoception and
79    exteroception emerges from the normative (formal) principles of active inference.
80    Furthermore, having an *in silico* subject at hand, means that we can simulate the effects of
81    various disconnections and pathophysiology. For example, we can examine the effect of
82    deafferentation of interoceptive signals on arousal, exteroceptive perception, and
83    (metacognitive) confidence placed in perceptual categorization. Indeed, we were able to go
84    beyond simulated deafferentation studies and ask what it would be like if we were able to
85    selectively lesion the precision of (i.e. confidence ascribed to) different sorts of beliefs; for
86    example, beliefs about 'what I am doing', beliefs about 'the state of the world', and beliefs
87    about 'the sorts of interoceptive and exteroceptive signals I expect to encounter'.

88    Second, we constructed our synthetic subject in such a way that the same paradigm
89    could be replicated in real subjects. The motivation for this is that the active inference
90    scheme used below has an associated process theory (Friston et al., 2017a). In other words,

3

91  neuronal and behavioral responses associated with inferential processes can be simulated

92  on a trial by trial basis. This means that we can use electrophysiological, eye tracking,

93  pupillometry and other physiological proxies to test various hypotheses that can be

94  instantiated in the model. Crucially, this provides a link between neuronal and behavioural

95  responses – as characterised by the latency between stimuli onset and autonomic responses

96  (e.g., heart rate acceleration or variability) or confidence judgements (i.e., responses to how

97  confident were you in your perceptual judgement?). In this paper, we will focus on the basic

98  phenomenology and (some counterintuitive) results. In subsequent work, we will use this

99  formalism to model real responses under various experimental manipulations.

100     In what follows, we briefly describe the generative model and inversion scheme used

101  to simulate cardiac arousal responses. We then demonstrate the results of anatomical

102  (deafferentation) lesions on perceptual and metacognitive behaviour, as well as simulated

103  belief updating. Finally, we will examine the effects on synthetic heart-rate variability when

104  changing the precision of various prior beliefs that underlie perceptual inference. We

105  conclude with a discussion of the implications for existing research in this area – and how

106  this research could be informed by a formal approach providing guidelines to discovery.

107

108

109  **Methods**

110

111  *Markov Decision Process*

112

113  The simulations reported below build upon the notion of active inference. This is a 'first-

114  principles' approach to understanding (Bayes) optimal behaviour. Simply put, active

115  inference treats the brain as using an internal (generative) model of the world to explain

116  exteroceptive, proprioceptive, and interoceptive sensory data. By optimizing beliefs about

117  variables in this model (perceptual inference), or by changing their internal or external

118    environment (action), creatures can ensure their sensations and prior beliefs are consistent[1].

119    A Markov decision process (MDP) is a form of probabilistic generative model that describes

120    the sequential dynamics of unobserved (hidden) variables (e.g., the current state of the

121    cardiac cycle) and the sensations they cause (e.g., baroreceptor signals). The hidden

122    variables of an MDP are hidden states ($s_\tau$) and sequences of actions or policies ($\pi$). The

123    generative model then embodies the conditional dependencies between these variables, as

124    expressed graphically in Figure 1. While we provide a brief overview here, we refer readers

125    to (Friston et al., 2017a) for more technical detail.

126

127    Hidden states generate observable sensory data with probabilities expressed in a likelihood

128    matrix **A**. The states evolve through time according to a transition probability matrix, **B** and

129    depend only on the state at the previous time, and on the policy, $\pi$,. Finally, we equip the

130    generative model with preferences (**C**), prior beliefs about initial states (**D**), and prior beliefs

131    about policies. Beliefs about policies have two parts. The first of these is a fixed bias (**E**). This

132    may be thought of as a habit; i.e., 'what I expect to do' *a priori*. The second is a belief that the

133    most probable policies are those that have the lowest expected free energy (**G**); i.e., 'what I

134    expect to do' after considering the consequences of action. A simple intuition for the latter is

135    to think of the selection between alternative courses of action as we might think of Bayesian

136    hypothesis testing (i.e. model comparison); namely, planning as inference (Attias, 2003;

137    Botvinick and Toussaint, 2012). Here, each policy can be thought of as an alternative

138    hypothesis about 'how I am going to behave'. These are evaluated in terms of prior beliefs

139    (**E**), and the (predicted) evidence future data affords (**G**). Just as free energy is used to

140    approximate the evidence data affords a hypothesis, expected free energy evaluates the

141    expected evidence, under beliefs about how data are actively generated. As expressed in

142    Figure 1, expected free energy can be separated into two parts. 'Risk' quantifies how far

143    predicted observations deviate from preferred outcomes. Minimizing this ensures

144    maintenance of homeostasis. 'Ambiguity' quantifies the uncertainty in the mapping from

---

[1] The term 'belief' here is used in the technical sense of a Bayesian belief, or probability distribution, typically considered to be sub-personal.

*Cardiac Active Inference*

145   states to outcomes. Minimizing this component ensures that salient, uncertainty-resolving

146   data are sought (leading to epistemic, information gathering, behavior).

147

148
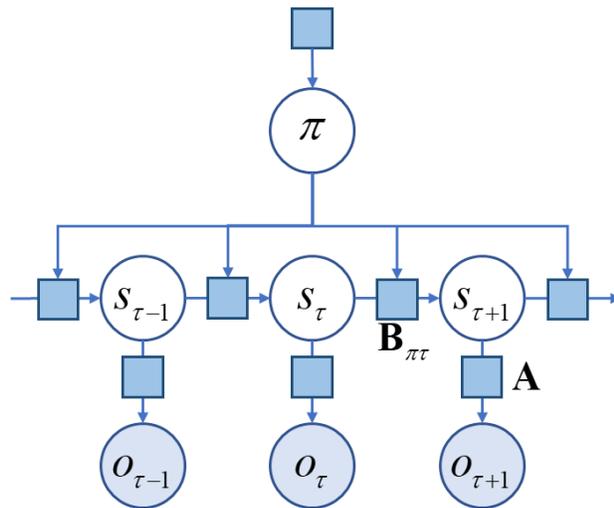
**Generative model**

$$P(\tilde{o}, \tilde{s}, \pi) = P(s_1)P(\pi)\prod_\tau P(o_\tau \mid s_\tau)P(s_{\tau+1} \mid s_\tau, \pi)$$

$$P(o_\tau \mid s_\tau) = Cat(\mathbf{A})$$
$$P(s_{\tau+1} \mid s_\tau, \pi) = Cat(\mathbf{B}_\pi)$$
$$P(o_\tau) = Cat(\mathbf{C})$$
$$P(s_1) = Cat(\mathbf{D})$$
$$P(\pi) = \sigma(\ln \mathbf{E} - \mathbf{G})$$

**Variational posterior**

$$Q(\tilde{s}, \pi) = Q(\pi)Q(\tilde{s} \mid \pi)$$
$$Q(s_\tau \mid \pi) = Cat(\mathbf{s}_\pi)$$

**Expected free energy**

$$\mathbf{G}_\pi = \underbrace{\mathbf{o}_{\pi\tau} \cdot (\ln \mathbf{o}_{\pi\tau} - \mathbf{C})}_{Risk} + \underbrace{\mathbf{H} \cdot \mathbf{s}_{\pi\tau}}_{Ambiguity}$$

149

150   **Figure 1**. A Markov decision process generative model: the factor graph *on the left* illustrates the

151   conditional dependencies, and independencies, between the variables in the generative model (see the main

152   text for a description of the variables). The variables are shown in circles (with filled circles showing observable

153   variables). An arrow from one variable to another indicates that the latter depends upon the former. The square

154   nodes each represent probability distributions. The panels *on the right* give the forms of the distributions

155   (associated with each square node) in the generative model, in addition to defining the expected free energy,

156   and specifying the factorization of the approximate posterior (variational) distributions the agent possesses.

157

158

159   *Synthetic Cardiac Arousal*

160

161   Using the MDP scheme detailed above, we set out to simulate a cardiac arousal

162   response to threatening stimuli (e.g., a vicious looking spider), in comparison to non-

163   arousing stimuli (e.g., some flowers). To do this, we had to define 'arousal' and its

164   interoceptive correlates. To keep things as simple as possible, we assumed the subject's

165   generative model included two sorts of hidden states (*interoceptive* and *exteroceptive* – and

6

166  that she could adopt two modes of engagement with the world (*relaxed* and *aroused*). These
167  sorts of generative models are generally cast as Markov decision processes, whereby
168  transitions among (hidden) states generate observable outcomes in one or more modalities.
169  The modalities considered here were *exteroceptive* (*non-arousing* versus *arousing* visual
170  stimuli) and interoceptive (the cardiac phase; *diastolic* or *systolic*). Having defined the nature
171  of the state space generating outcomes, this model can then be parameterised in a relatively
172  straightforward fashion as outlined above. For any set of **A**,**B**,**C**,**D,** and **E** parameters, one can
173  then simulate active inference using standard marginal message passing schemes (Parr et
174  al., 2019) to optimize expectations about hidden states of the world – and the action or policy
175  currently being pursued (technically, a policy is a sequence of actions. In what follows, we
176  only consider policies with one action) (Friston et al., 2017a, 2017c).

177      Crucially, inference about policies rest upon prior beliefs that the policies will
178  minimise expected free energy in the future. This expected free energy has both epistemic
179  and instrumental terms; namely; the ability of any particular course of action to resolve
180  uncertainty about hidden states (known as salience, Bayesian surprise, information gain,
181  *etc.*) (Barto et al., 2013; Itti and Baldi, 2009; Oudeyer and Kaplan, 2009; Schmidhuber, 2010)
182  and the pragmatic affordance (known as expected value, utility, reward, *etc.*) as specified by
183  the prior preferences (Friston et al., 2015).
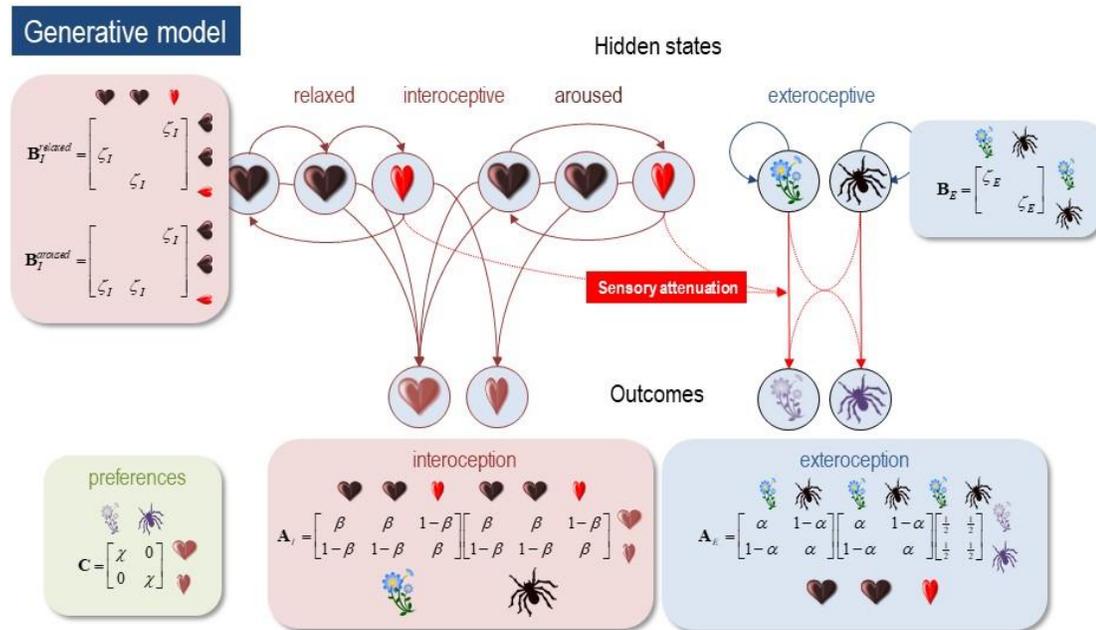
**Figure 2**: *the generative model*. This schematic illustrates how hidden states cause each other and sensory outcomes in the interoceptive and exteroceptive domain. The upper row describes the probability transitions among hidden states, while the lower row specifies the outcomes that would be generated by combinations of hidden states that are inferred on the basis of outcomes. The green panel specifies the models prior preferences; namely, the sorts of outcomes it expects to encounter. Please see main text for a full explanation. Although this figure portrays interoceptive and exteroceptive outcomes as separate modalities, they were in fact modelled as combinations – so that the prior preferences could be evaluated (this is necessary because the preferred physiological outcome depends upon the visual cue). In this model, the precisions are denoted by Greek letters and control the fidelity of various probabilistic mapping is (i.e., the likelihood or **A** matrices and the transition or **B** matrices).

To capture the fundaments of multimodal integration – of interoceptive and exteroceptive modalities – we assumed the following, reasonably plausible, form for the model. The synthetic subject had to infer which of two policies she was pursuing: a *relaxed* policy or an *aroused* policy. These are defined operationally in terms of transitions among interoceptive states. Here, we model this in terms of two distinct forms of cardiac cycling among *diastolic* and *systolic* bodily states. When *relaxed*, the probability transitions among cardiac states

8

202    meant that there were two phases of *diastole* and one of *systole*. Conversely, when *aroused*,

203    the first *diastolic* state jumped immediately to *systole*. In brief, this means that being aroused

204    causes cardiac acceleration and the average amount of time spent in systole. The outcomes

205    are generated by these states were isomorphic; in other words, there was a simple likelihood

206    mapping from states to sensations; such that the subject received a precise or imprecise

207    interoceptive cue about the current cardiac status (i.e., *diastole* or *systole*).

208            On the exteroceptive side, we just considered two states of the visual world; namely,

209    the subject was confronting an *arousing* or *non-arousing* visual object. The corresponding

210    visual modality again had two levels (*arousing* versus *non-arousing* picture). Crucially, the

211    fidelity or precision of this mapping depended upon the interoceptive state. When the

212    subject was in *systole*, this mapping became very imprecise. In other words, all outcomes

213    were equally plausible under each hidden state of the visual world. Conversely, during

214    *diastole*, there was a relatively precise likelihood mapping. This is the crucial part of our

215    model that links the state of the body to the way that it samples the world. Put simply, precise

216    visual information is only available during certain parts of the cardiac cycle, which itself

217    depends upon the state of arousal (i.e., the policy currently inferred and selected). This can

218    be thought of as a simple approximation of cardiac and other bodily timing effects, expressed

219    as a momentary occlusion or attenuation of sensory input by (for example) afferent

220    inhibitory baroreceptor effects (Bonvallet and Bloch, 1961; Lacey and Lacey, 1978), or by

221    the brief flooding of the retina during cardiac contraction.

222            This simple structure produced some remarkable results that speak to the intimate

223    relationship between interoception and exteroception. These phenomena (see below) rest

224    upon the final set of beliefs; namely, preferred outcomes. Here, the subject believed that she

225    would be, on average, in a *systolic* state when confronted with an arousing picture and in a

226    *diastolic* state otherwise. These minimal prior preferences then present the subject with an

227    interesting problem. She has to choose between extending periods of precise evidence

228    accumulation (i.e., a *relaxed* state with more *diastolic* episodes) and sacrificing precise

229    information, via cardio-acceleration, should she infer there is something arousing 'out there'.

230    However, to infer what is 'out there', she has to resolve her uncertainty, through epistemic

231    foraging; i.e., maintaining a *relaxed* state. We therefore hypothesised that at the beginning of

232   each trial or exposure to a picture[2], subjects would be preferentially in a relaxed state until

233   they had accumulated sufficient evidence to confidently infer the visual object was *arousing*

234   or not. If *arousing*, she would then infer herself to be aroused and enter into a period of

235   cardio-acceleration (illustrated in Figure 3).

236        By carefully adjusting the precision of sensory evidence (through adjusting the **A**

237   matrix), we could trade-off the evidence accumulation against these imperatives to simulate

238   the elaboration of an arousing response to, and only to, *arousing* stimuli. Furthermore, we

239   anticipated that a failure to implement a selected policy of arousal would both confound

240   inference about the policy being pursued (i.e., an aroused state of mind) and – importantly –

241   confidence about the exteroceptive state of affairs. The latter can be measured quantitatively

242   in terms of the entropy or average uncertainty over hidden exteroceptive states (after taking

243   a Bayesian model average over policies). This leads to the prediction that confidence in

244   perceptual categorisation would not only evolve over time but would depend upon

245   interoceptive inference. We tested this hypothesis *in silico* through various lesion

246   experiments reported in the subsequent sections (Figures 3 - 5). In what follows, we

247   illustrate the belief updating and arousal responses under 'normal' priors (i.e. precisions)

248   based upon the generative model above (summarized graphically in Figure 2).

249   *Simulations*

250   We implemented a minimal model of interoceptive and emotional inference – in the sense

251   that one's state of active engagement with the world may be inferred from its interoceptive

252   and exteroceptive consequences. In this minimal model, the two domains of perception are

253   coupled by – and only by – sensory attenuation: i.e., attenuation of sensory precision in the

254   visual domain during (inferred) systole. Precision refers to the reliability or confidence

255   ascribed to a given probabilistic belief. Within this model there are four kinds of precision;

256   namely, sensory precision in the visual ($\alpha$) and cardiac ($\beta$) domains and the precision of state

257   transitions among interoceptive and exteroceptive states. For clarity, we will refer to the

258   precision of transitions as (inverse) *volatility* and use *precision* to refer to sensory (i.e.

---

[2] Every simulation started off with a weak prior over hidden states that the picture was not arousing – and a weaker prior in favor of the *relaxed* policy.

10

259    likelihood) precision. In this example, because there are only two states, the corresponding

260    parameters of the generative model control both the expected contingencies and their

261    precision. In other words, when α (or β) decreases to 1/2, sensory signals become imprecise

262    and completely ambiguous. In what follows, we will focus on manipulations of precision

263    under a canonical volatility of $\zeta = 0.9$. In other words, we will assume that our synthetic

264    subject believes state transitions among phases of the cardiac cycle follow each other fairly

265    reliably with a 90% probability. Similarly, if there is a flower 'out there', then there is a 90%

266    probability that it will remain there at the next sample. Cardiac and visual stimuli were

267    generated by the same precisions and volatilities as assumed by the subject's generative

268    model.

269    We conducted three sets of simulations to illustrate the sorts of behaviours that

270    emerge under this active inference scheme – and to establish the construct validity of the

271    model in relation to empirical phenomena that speak to the influence of interoception on

272    exteroception and *vice versa*. This enabled us to illustrate the basic phenomenology of our

273    agent – in terms of simulated perceptual inference and cardiac physiology – under some

274    differing levels of sensory precision.

275
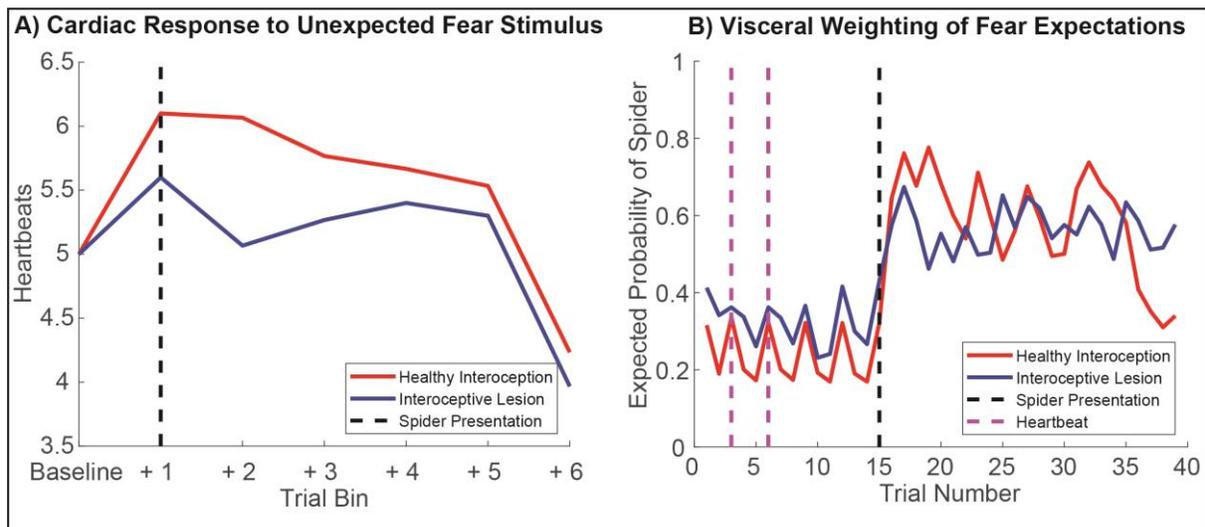


277    **Figure 3**. **Simulated Physiology and Perceptual Inference**. To establish the face validity of our

278    model, we first set out to reproduce some basic psychophysiological phenomenology and establish how these

279    phenomena change under 'healthy' (i.e., normative) versus 'visceral lesion' parameter settings. To do so, we fed

280    agents a fixed sequence of cardiac and exteroceptive stimuli, such that the first 14 trials constituted a 'baseline'

281    period of cardiac quescience (i.e., a steady heart rate), in the absence of arousing stimuli. On the 15th trial, an

11

282    unexpected arousing stimulus (a 'spider') is presented and a further 85 trials simulated. This simulation was

283    repeated for 60 simulated participants, each with randomized starting values, half of which had 'lesioned'

284    interoceptive precision (β = 0.5, blue lines). Under these conditions our synthetic subjects exhibit a clear

285    'startle' or 'defense' reflex (Graham and Clifton, 1966; Sokolov, 1963), characterized by an immediate cardio-

286    acceleration (left panel) and a dramatic shift in the posterior expectation of encountering another threatening

287    stimulus. Interestingly, during the baseline period the posterior expectation of encountering a threat stimulus

288    oscillates with the heartbeat; i.e., the lesioned subjects show both an attenuation of the cardiac response and a

289    blunted belief update. Note that for the right panel, only trials 1-40 are shown. On the left, blue lines show

290    summed heartbeats (time spent in systole) for 15-trial bins; on the right, lines depict the median posterior

291    probability that the agent will see a spider on the next trial. See *Methods* and *Results* for more details.

292

293    In the first set of simulations (Fig. 3), we focused on the physiological and psychological

294    response to arousing stimuli. To do so, we tested the hypothesis that the unexpected

295    presentation of a 'spider' would induce an aroused state – as reflected in an increased heart

296    rate – and a greater posterior expectation of encountering an arousing spider stimulus on

297    the next trial. To evaluate this hypothesis, we supplied the subject with a fixed sequence of

298    15 stimuli – in both the cardiac and visual domains – and examined the posterior beliefs

299    about the next exteroceptive state following a period of relaxed cardiac input. Note that this

300    is possible precisely because our generative model includes beliefs about the future –

301    including the next hidden state and subsequent sensory sample. Here, we used as outcome

302    measures the agent's evoked cardiac acceleration response (calculated by binning the

303    number of siastole events across the experiment) and the agent's posterior belief that the

304    next stimulus would be threatening. These simulations were repeated 60 times with

305    randomized starting values, such that the first thirty 'healthy' agents where compared to an

306    'interoceptive lesion' group for whom interoceptive precision had been attenuated (β = 0.5).

307    This enabled us to not only establish the interaction of fear expectations and cardiac arousal,

308    but also to demonstrate how these responses change when interoceptive sensory precision

309    is ablated.

310    In the second set of simulations (Fig. 4), our focus moved from perceptual to

311    metacognitive inference. Here, we examined the interaction between exteroceptive and

312    interoceptive sensory precision on the one hand and their coupling to cardiac timing and

313    metacognition (posterior confidence) on the other. Our goal here was to illustrate how both

314    interoceptive and exteroceptive precision interact to influence metacognitive inference, and

315    to link these to empirical findings showing that cardiac arousal biases metacognition (Allen

316    et al., 2016b; Hauser et al., 2017a). For these, we used the uncertainty about inferred

317    exteroceptive and interoceptive states (as quantified by the summed entropy of posterior

318    beliefs for each state) as outcome measures, simulated under a range of cardiac and visual

319    precision settings (figure 3A). To further illustrate how these effects oscillate with the

320    cardiac rhythm, we separated these measures for each phase of the cardiac cycle (early

321    diastole, late diastole, systole). We then repeated these analyses comparing 'healthy'

322    interoceptive inference agents ($\alpha$ & $\beta$ = 0.9), to agents for whom either exteroceptive or

323    interoceptive precision was lesioned ($\alpha$ or $\beta$ = 0.5, respectively). In virtue of our coupling of

324    exteroceptive sensory precision to the cardiac cycle, we anticipated that metacognitive

325    confidence (outscored by the negative entropy of posterior beliefs) would depend on the

326    precision of both interoceptive and exteroceptive states, and that this effect would clearly

327    oscillate with the cardiac cycle. Further, we expected in the extreme case of our 'lesioned'

328    subjects, these effects would be further exacerbated such that interoceptive and

329    exteroceptive uncertainty would increase dramatically, under their respective lesion

330    conditions.

331
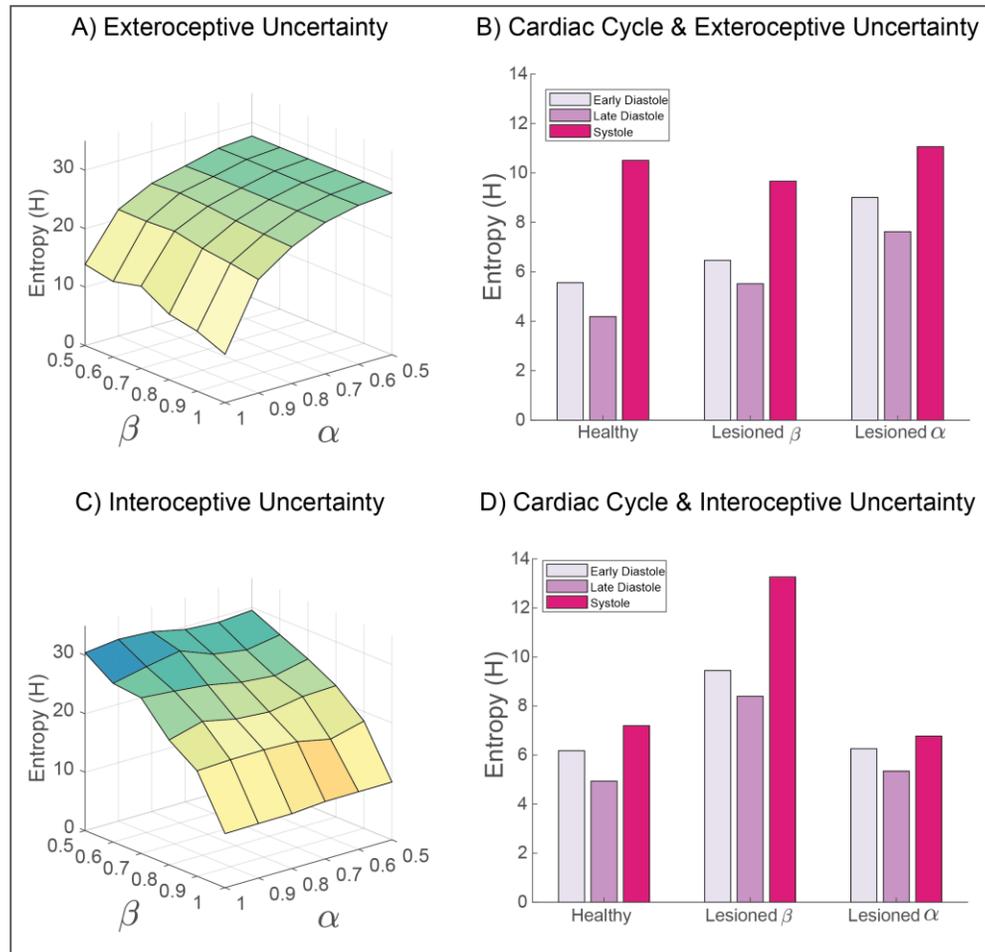
**Figure 4**. Simulating the influence of interoceptive and exteroceptive precision on metacognitive uncertainty. To explore how interoceptive inference influences metacognition, we measured the summed entropy of beliefs for both exteroceptive (top panels) and interoceptive (bottom panels) states. By simulating the full range of sensory precision values, from lesioned precision ($\alpha$ or $\beta$ = 0.5) to 'hyper-precision' ($\alpha$ or $\beta$ = 1), the predominant pattern of interactions is revealed. **A)** For exteroceptive inferences (i.e., the agent's belief that a spider or flower is present), the principle entropy gradient is characterized by reductions in exteroceptive precision. This effect is modulated in part by interoceptive precision; for example, the lowest uncertainty is obtained when interoceptive and exteroceptive precision are maximal. **B)** Separating exteroceptive uncertainty by each phase of the cardiac cycle reveals a clear effect of the heartbeat on belief entropy, which is modulated most strongly by lesioning the precision of exteroceptive predictions. Lesioning interoceptive uncertainty does raise the overall level of exteroceptive uncertainty, but to a lesser degree. Note that altering exteroceptive precision only affects the diastolic phases (as precision is already attenuated during systole). Interoceptive lesions preclude precise inferences about the cardiac phase, so reduce the discrepancy in uncertainty between these phases. **C)** Similar to exteroceptive belief, interoceptive metacognition is predominately influenced by interoceptive precision. **D)** The cardiac cycle also modulates the overall

14

348    uncertainty of interoceptive beliefs; this effect is greatly increased when interoceptive precision is lesioned.

349    Interestingly, exteroceptive lesions primarily reduce the differentiation between cardiac states.

350

351    Finally, to complement these simulations we modelled the response of first and second order

352    statistics of the physiological responses to changes in sensory precision. These were based

353    upon simulated heart rate (frequency of systole) and the heart rate variability (HRV)

354    assessed over multiple trials or heartbeats (Fig. 5). Our objectives here were; 1) to test the

355    hypothesis that fluctuations in both low-and high- frequency synthetic heart rate variability

356    can be produced by altering the balance of interoceptive sensory precision versus the prior

357    precision for the aroused sympathetic policy, and 2) to illustrate how generative modelling

358    of interoceptive active inference can be used to phenotype maladaptive inference

359    parameters from observed heart-rate data (i.e., interoceptive inference phenotyping). For

360    this analysis, we simulated 1000 trials under three canonical parameter settings designed to

361    resemble potential neuropsychiatric phenotypes of interest: healthy interoception ($\alpha$ = 0.8,

362    $\beta$ = 0.8, prior probability of parasympathetic policy = 55%), hyper-precise interoceptive

363    sensation ($\alpha$ = 0.8, $\beta$ = 1, prior probability of parasympathetic policy = 55%), and hyper-

364    precise arousal priors ($\alpha$ = 0.8, $\beta$ = 0.8, prior probability of sympathetic policy = 75%).

365         The resulting time-series of systole events from each agent were then convolved with

366    a canonical QRS-wave response function and transformed into normalized beat-to-beat RR-

367    intervals. To normalize the (arbitrary) sampling rate of each time-series, we assigned a

368    350ms repetition time (TR) for each state of the MDP simulation, such that the healthy agent

369    had a heart rate of approximately 60 BPM. The time intervals between successive synthetic

370    R-peaks was then calculated. As the RR interval data is unevenly sampled, the time series

371    was linearly interpolated. The power spectrum was then estimated using Welch's method.

372    In line with conventional HRV analysis, the power spectra were then categorized into four

373    frequency bands corresponding to ultra-low (0 - 0.04 Hz), low (0.04- 0.15 Hz), high (0.15 -

374    0.4 Hz), and very-high (> 0.4 Hz) frequency categories. Finally, to summarize the

375    physiological response of each agent, we calculated the beats per minute (BPM) and the ratio

376    between low and high frequency components (LF/HF), i.e., sympathovagal gain or balance.

377    Sympathovagal valance is thought to index the balance of sympathetic and vagal outflows

378    and is frequently interpreted implicated in stress and other psychophysiological and clinical

379    disorders (Malliani et al., 1991; Strigo and Craig, 2016) but see (Eckberg Dwain L., 1997;

380    Heathers, 2012) for critique. Sympathovagal balance was calculated as the ratio of area

381    under the curve (AUC) for low and high-frequency HRV; $AUC^{LF}/AUC^{HF}$.

382
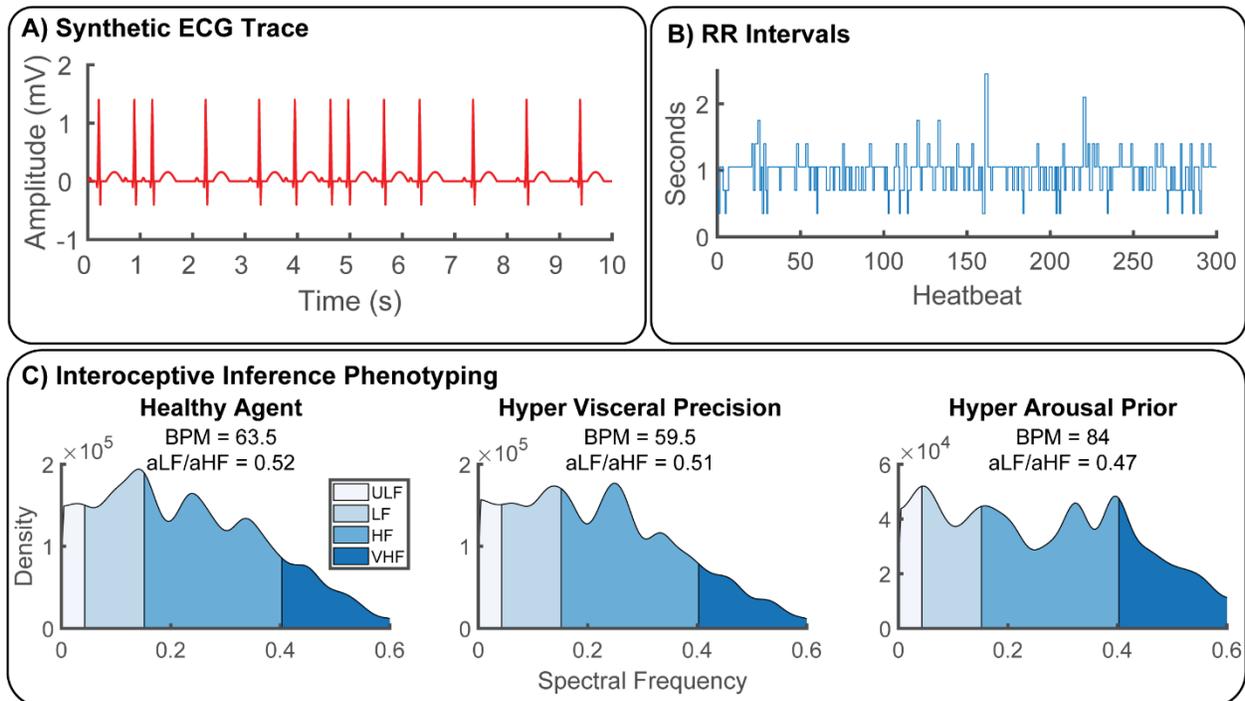
383



384    **Figure 5.** Synthetic Heart-Rate Variability (HRV) and Interoceptive Computational Phenotyping. To

385    illustrate the potential of our approach as a generative model of physiological reactivity, we produced synthetic

386    heartbeat traces and analyzed these with a standard time-frequency approach under various canonical

387    parameter settings. **A)** Synthetic ECG traces produced by convolving a standard QRS-wave function with systole

388    events generated by our model. **B)** These where then transformed into RR-intervals by assuming an 350ms

389    sampling rate, **C)** Power spectra of RR-intervals were calculated using Welch's method and categorized as ultra-

390    low (ULF), low, (LF), high (HF), and very high frequency (VHF) bands for each simulated agent. Physiological

391    responses were then summarized in terms of beats-per-minute (BMP) and sympathovagal balance (ratio of

392    area under curve for each frequency band, aLF/aHF) (Malliani et al., 1991). To illustrate the potential of our

393    approach for interoceptive computational phenotyping, we simulated three different agents – one with healthy

394    interoceptive inference (bottom left), another with hyper-precise visceral sensations (bottom middle), and

395    another with hyper-precise priors for the aroused (sympathetic) policy (bottom right). These each produce

396    unique interoceptive inference 'fingerprints'; i.e., the individual patterns of heart-rate variability produced by

397    these parameter settings. In this example, hyper-precise visceral sensations reduce heart-rate and shift overall

398    peak frequency to the high-frequency domain, whereas hyper strong arousal priors induce strong heart-rate

399  acceleration coupled with attenuated ultra-low and ultra-fast oscillations. In the future, these idiosyncratic

400  patterns could be used to identify maladaptive interoceptive inference from heart-rate data.

401

402

403  **Results**

404

405  *Simulated Physiology and Perceptual Active Inference.*

406  To establish the face validity of our model, we simulated the basic psychophysiological

407  behavior of our active inference agent. This involved simulating a fixed series of stimuli

408  (states) in which the heartbeat was forced to remain relaxed – and only non-arousing

409  (flower) stimuli were presented. On the 15th trial, an unexpected spider stimulus was

410  presented, and the simulation continued for a further 85 trials. Thus, by evaluating the

411  evolution of the agent's synthetic interoceptive physiology and exteroceptive beliefs, before

412  and after the quiescent baseline period, we hoped to reproduce and illuminate well-known

413  psychophysiological phenomenon such as the defensive startle reflex (Graham and Clifton,

414  1966; Sokolov, 1963).

415      This analysis, illustrated in Figure 3, revealed several interesting aspects of

416  interoceptive active inference. Over 60 simulations there was a clear and robust increase in

417  heart-rate acceleration, following the presentation of the unexpected or novel threat

418  stimulus. During subsequent experiences of its own heartbeat and spiders or flowers, this

419  response habituates, resulting in a gradual heart-rate deceleration from the evoked cardiac

420  response. This robust modulation of heart-rate was accompanied by a jump from an

421  expected probability of encountering a spider of about 25% to almost 65% following the

422  spider presentation. This combined response of both the heartbeat and fear-expectations is

423  further underscored by the curious oscillation of cardiac states and the expected probability

424  of observing a spider; note the uptick in expectations of approximately 5% on each systole

425  event (denoted by the pink dotted line on Figure 3, right panel). A simple explanation for

426  this result is that, during presentation of a stream of flowers, we can confidently infer a safe

427  external environment. This accounts for the relatively low probability of spiders in the

428  earlier part of the plot. However, during systole, attenuated integration of exteroceptive data

17

429    leads to greater uncertainty. Going from a confident inference in the absence of a spider to a

430    more uncertain inference necessarily increases the probability of a scary environment

431    during this cardiac phase. This offers a simple perspective on previous experimental work

432    suggesting that fear-stimuli are potentiated when presented in synchrony with the heart

433    (Garfinkel et al., 2014; Garfinkel and Critchley, 2016); namely, that a mechanism underlying

434    this effect can be found in the link between cardiac active inference and fear expectations. In

435    short, under generative models of an embodied world – in which sensory sampling depends

436    upon fast fluctuations in bodily states – there is a necessary dependency of Bayesian belief

437    updating (i.e., perceptual inference) across all modalities on introception.

438        When comparing these effects in the healthy agent to our sample of 'lesion patients',

439    a few sensible but counter-intuitive consequences ensue. In the physiological domain, when

440    presented with the unexpected arousal stimulus, the lesioned agent shows a blunted cardiac

441    acceleration response, which remains diminished throughout the simulated trials. This

442    blunting effect is mirrored for fear expectations in the immediate post-stimulus (e.g., trials

443    15-20) period, further underlining the close link between visceral and exteroceptive

444    inference in our agent. The reason for this blunting likely results from the differing

445    exteroceptive precision anticipated during different cardiac phases (see also Fig. 4B). A

446    visual impression – consistent with a spider – is highly informative during diastole but must

447    be treated with suspicion during the sensory-attenuated systolic phase. This implies a

448    blunting of belief-updating in response to a spider, when we are unsure of cardiac phase

449    (compared to when we are confident of a diastolic phase).  A further interesting result is

450    found when examining the controlled baseline period (trials 0-15); baseline fear

451    expectations in the interoceptive lesion group are actually slightly enhanced by about 5-10%

452    posterior probability. This lends an interesting embodied twist to the literature on 'circular

453    inference', psychosis and hallucinations (Denève and Jardri, 2016; Powers et al., 2017),

454    suggesting that the disruption of interoceptive precision may be one mechanism underlying

455    hallucinations, particularly those that are affective and/or somatic in nature.

18

456    *Simulating the Influence of Sensory Precision on Metacognition*

457    We next performed a series of simulations to tease apart how interoceptive and
458    exteroceptive precision (and their disruption) influence 'metacognition'; that is the
459    uncertainty in our agent's beliefs. To do so, we first measured the Shannon entropy for
460    interoceptive and exteroceptive inferences (summed across both factors of posterior beliefs)
461    under a full range of precision settings from 0.5 - 1. To highlight the oscillatory nature of
462    cardiac effects, we then calculated the same entropy measure separately for each cardiac
463    state (early diastole, late diastole, systole). Finally, we compared these 'healthy' simulations
464    to extreme degradations in sensory precision (exteroceptive and interoceptive 'lesions'), to
465    better understand how disruptions of each modality are integrated in metacognition.

466    This analysis revealed first of all that, in our simplified model, metacognitive
467    uncertainty is largely influenced by the unimodal precision of each domain. For both
468    exteroceptive and interoceptive inferences, the slope of the uncertainty gradient (Fig. 4A &
469    C) was predominantly characterized by degradations in the precision of the corresponding
470    modality. However, this modularity is not complete; exteroceptive uncertainty is at its lowest
471    when interoceptive and exteroceptive precision are maximal. Similarly, although
472    interoceptive uncertainty is largely driven by interoceptive precision, small interactions
473    with exteroceptive precision can be observed in the plotted uncertainty gradient. One
474    interesting isomorphism, however, is that overall interoceptive uncertainty is less affected
475    by exteroceptive precision. This is likely due to that fact that in our model, the cardiac cycle
476    directly modulates exteroceptive precision, whereas exteroceptive states only indirectly
477    modulate interoceptive responses, via policy selection.

478    This intricate relationship of the cardiac cycle and metacognitive uncertainty is
479    further teased apart in Figure 4B, which shows clearly that exteroceptive confidence
480    oscillates with each phase of the heartbeat, being highest at diastole. This is an unsurprising
481    feature of our model: on each diastole, phase exteroceptive sensory precision drops
482    effectively to null. Interestingly however, average exteroceptive uncertainty is modulated in
483    a fairly linear fashion by visceral and exteroceptive lesions: average entropy is increased
484    modestly by lesioning interoceptive precision and more robustly by exteroceptive lesions.
485    Whereas interoceptive lesions caused the greatest increase in interoceptive entropy,

486 exteroceptive lesions seem to exert a specific effect of unbinding entropy from the individual

487 cardiac state, again mirroring the isomorphic representation of these states in uncertainty.

488 This is a sensible finding, as the manipulation leads to relatively high uncertainty in the

489 mapping between hidden states and outcomes during all cardiac phases, not just during the

490 previously attenuated systolic phase. This sort of chronic hypo-arousal – as a consequence

491 of a failure to contextually modulate precision – is not unlike that which may underwrite the

492 negative symptoms of schizophrenia or depression.

493 *Synthetic Heart-Rate Variability (HRV) and Embodied Computational Phenotyping*

494 In our final set of simulations, we illustrated how the interoceptive inference approach

495 developed here offers a new means for analyzing and interpreting fluctuations in observed

496 physiological data. Our goal here was to demonstrate the potential for generative modelling

497 and 'embodied computational phenotyping'; i.e., the identification of specific parameters of

498 brain-body interaction underlying maladaptive interoceptive inference in psychiatric and

499 other health-harming disorders; e.g., (Peters et al., 2017).

500 To this end, we generated synthetic cardiac data by convolving our train of cardiac

501 events with an ECG response waveform. Following standard methods, we then calculated the

502 normalized beat-to-beat intervals and performed a time-frequency analysis of the resulting

503 RR-interval data. By repeating this analysis for a 'healthy' agent under normative values, an

504 agent with interoceptive 'hyper-precision' (i.e., $\beta = 1$), and an agent with an overly precise

505 prior beliefs about its own arousal, we illustrate how individual HRV fingerprints are linked

506 to unique patterns of interoceptive active inference.

507 This analysis showed that, despite the exceedingly simple (biomechanically speaking)

508 conditions of our model, sensible and interesting patterns of heart-rate variability emerge

509 for different combinations of interoceptive sensory and prior precision. Specifically, we

510 found that whereas the healthy agent exhibited a relatively relaxed profile in terms of heart

511 rate and sympathovagal balance (BMP = 63.5, peak frequency = 0.14 Hz) – predominated by

512 low versus high frequency oscillations (aLF/aHF = 0.52) – an agent with hyper-precise

513 visceral sensations exhibited a mild downshift in heart-rate coupled (BPM= 59.5) with an

514 overall increase in high-frequency oscillations (aLF/aHF = 0.51, peak frequency = 0.25). In

515  contrast, the agent with hyper-precise arousal priors showed a strong bimodal modulation

516  of both ultra-low and ultra-high frequencies HRV (peak frequencies = 0.04 Hz & 0.34 Hz,

517  respectively), coupled with a strong increase in heart-rate (BPM = 84) and high versus low-

518  frequency outflow (aLF/aHF = 0.47). These results speak to the unique role of different

519  active inference parameters in producing highly idiosyncratic patterns of HRV variability. In

520  the future, our model may be enhanced to subserve computational phenotyping of individual

521  differences and/or patient subgroups categorized by the balance of visceral precision and

522  arousal policy priors from raw HRV data alone.

523

524

525  **Discussion and Conclusions**

526

527  In the present work, we have introduced the first formal model of interoceptive inference as

528  applied to emotion, exteroceptive perception, and metacognitive uncertainty. Through a

529  variety of simulations, we demonstrated that this model can reproduce a variety of

530  psychological and physiological phenomena, each of which speak to a unique domain of the

531  burgeoning interoceptive inference literature (Allen and Friston, 2018; Feldman and Friston,

532  2010; Seth, 2013), and the application of interoceptive inference to computational

533  psychiatry (Owens et al., 2018; Petzschner et al., 2017). This formulation of interoceptive

534  inference reproduces some of the finer details of physiological responses to arousing stimuli

535  that, crucially, are emergent properties under the simple assumption that people use

536  generative models to infer the state of their lived world.

537      The form of the generative model and (neurobiological implausible) belief updating

538  used in this paper are generic: exactly the same scheme has been used to simulate a whole

539  range of processes, from neuroeconomic games to scene construction and attentional neglect

540  (Friston et al., 2017a; Parr and Friston, 2018). The key aspect of the generative model

541  introduced here is that the quality (i.e., precision) of sensory information depends upon

542  fluctuations in (inferred) autonomic states. This simple fact underwrites all of the

543  phenomenology illustrated above; both in terms of simulated physiology and accompanying

544  belief updates. The explicit inclusion of interoception into active inference licenses us to talk

545  about 'fear' and in the sense that affective inference is thought to emerge under models that

546 generate multimodal predictions that encompass the interoceptive domain. Furthermore,

547 casting everything as inference enables a metacognitive stance on belief updating, in the

548 sense that one can quantify uncertainty invested in beliefs about states of the body, states of

549 the world and, indeed, states of (autonomic) action.

550 In particular, we show that by simulating periodic attenuation of exteroceptive

551 sensory inputs by the cardiac cycle, affective expectations become intrinsically linked to

552 afferent interoceptive signals through a startle reflex-like phenomenon. This linkage not only

553 induces oscillatory synchrony between the heartbeat and exteroceptive behavior, but also

554 propagates to metacognitive uncertainty (i.e., the entropy of posterior beliefs). This latter

555 finding speaks to numerous reports of metacognitive bias (e.g., confidence-accuracy

556 dissociation) by illustrating how the precision of interoceptive states can directly influence

557 exteroceptive uncertainty (Allen et al., 2016b; Boldt et al., 2017; Spence et al., 2016). By

558 simulating synthetic heart-rate variability (HRV) responses, we further illustrated how

559 idiosyncratic patterns of aberrant interoceptive precision-weighting can be recovered

560 through generative modelling of physiological responses, opening the door to computational

561 phenotyping of disordered brain-body interaction in the spirit of (Schwartenbeck P and K

562 Friston 2016). In what follows, we outline some of what we view as the most promising

563 future directions for this work, sketch a proposed neuroanatomy underlying our model, and

564 point out a few limitations for consideration.

565 By focusing on the periodic nature of the cardiac cycle, and concomitant influences

566 on exteroceptive perception, our goal was to provide an initial proof-of-principle, illustrating

567 how visceral and exteroceptive signals may be combined under active inference. Our aim

568 was not to suggest that our model provides the ultimate view of interoceptive inference;

569 indeed, we view the present work as a starting point that can be taken forward in a variety

570 of research directions, some of which are outline below.

571 In this paper, we formalized the hypothesis that frequently reported effects of cardiac

572 timing on perception could arise as a function of periodic sensory attenuation – but the

573 reader should feel encouraged to test their own hypotheses within the openly available MDP

574 framework. Our intention here was also not to prioritize cardiac-brain interaction over e.g.,

575 gastric or respiratory cycles, but instead to provide a toy example, to show how these

576 systems may be subjected to formal analyses. This was motivated by the large predominance

22

577    of research on cardiac-brain interaction; however, we do anticipate that the periodic
578    attenuation of sensory precision by visceral signals is likely to provide a general explanation
579    of brain-body interaction.

580        Neurophysiologically, the principal means by which cardiac signals influence the
581    central nervous system is through the afferent cardiac baroreceptors. These pressure-
582    sensitive neurons, located primarily in the aorta and carotid artery, are triggered by the
583    systolic pressure wave generated when the heart contracts. Far from being restricted to
584    homeostatic function only, it was first reported (nearly a century ago) that afferent
585    baroreceptor outputs induce a general inhibitory effect on cortical processing (Bonvallet et
586    al., 1954; Bonvallet and Bloch, 1961; Koch, 1932). These findings were later extended by
587    Lacey and Lacey (1978) who proposed the "neurovisceral afferent integration hypothesis",
588    positing that cardiac acceleration and deceleration serve to respectively disengage or engage
589    with an exteroceptive stimulus via cortical inhibition.

590        In parallel, the soviet psychologist Evgeny Sokolov proposed that novelty (but not
591    threat) evoked heart-rate deceleration was a core component of the 'orienting reflex'
592    (Sokolov, 1963). By reducing overall cardiac output, this reflex served to limit the
593    contribution of cardiac signals to cortical noise boosting overall signal-to-noise ratio[3]. In
594    contrast, Sokolov theorized that the defensive startle reflex – in which an extremely strong
595    (e.g., the loud bang of a starting gun) or unexpectedly aversive (e.g., the sudden presentation
596    of a spider) stimulus evokes cardiac acceleration – facilitated the disengagement of cortical
597    processing, to initiate fight-or-flight responses. These theories in turn sparked a wave of
598    empirical studies attempting to link cardio-acceleration and deceleration responses to
599    increased or decreased exteroceptive sensitivity, which continues to this day (Azevedo et al.,
600    2017; Cohen et al., 1980; Delfini and Campos, 1972; Edwards et al., 2009; Elliott, 1972;
601    Garfinkel et al., 2014; Ghione, 1996; Park et al., 2014; Salomon et al., 2016; Sandman et al.,
602    1977; Saxon, 1970; Velden and Juris, 1975).

---

[3] Sokolov (1963) described the orienting reflex as an 'embodied' mechanism for boosting to signal-to-noise and thus enhancing processing of the oddball stimulus. The reflex consists primarily of the rapid deployment of saccades to the oddball stimulus, freezing of the muscles of the head and neck so as to orient the visual organs towards the stimulus, and an immediate cardiac deceleration. In light of their inhibitory influence, the cardiac deceleration was thought to primarily reduce cortical noise; when coupled with the other bodily components of the response it was thought that effective overall signal would be maximized.

603       While these findings highlight the intricate relationship between cardiac timing and

604    exteroceptive psychophysics, so far a consistent pattern of findings (e.g., sensory signal

605    enhancement and/or inhibition) has failed to emerge (see Elliott, 1972 for one critique). A

606    cursory review of this literature reveals evidence for both exteroceptive enhancement and

607    suppression, depending upon the specific nature of the exteroceptive stimuli (i.e., whether

608    they are inherently aversive, sociocultural, or neutral in nature), the context of the arousal

609    (including specific stimulus and response timing), and other psychophysiological

610    moderators; such as age, gender, and overall physical fitness. Accordingly, more recent

611    proposals have focused on more modality-specific exteroceptive enhancement by cardiac

612    signals. For example, that cardiac-exteroceptive effects specifically potentiate fear or threat

613    signals (Garfinkel and Critchley, 2016) or the generation of a subjective first-person

614    viewpoint (Park and Tallon-Baudry, 2014).

615       We offer a unique synthesis of these views, expressed in terms of interoceptive

616    inference. In our model, the cyclic influence of the heart on exteroception is exerted primarily

617    through the attenuation of sensory precision on each systolic contraction, which in turns

618    influences the selected (multimodal) arousal policy as determined by the agent's

619    preferences. The coupling of sensory attenuation to the cardiac cycle endorses the notion

620    that baroreceptors exert an inhibitory influence on the brain. Beyond this direct effect, our

621    model can also be understood in light of the well-known relationship between intrinsic noise

622    fluctuations in the brain and cardio-respiratory cycles (Birn, 2012; Karavaev et al., 2018).

623    Physiological oscillations exert non-neuronal influences on spontaneous brain activity via a

624    variety of more or less direct causal influences; for example, at each heart beat visual input

625    to the retina is briefly attenuated by a pulsatile blood inflow. Similarly, with each cardio-

626    respiratory cycle, fluctuations in cerebral pulsatile motion and blood pressure induce

627    neurons to spontaneously fire, shaping the 'infraslow' brain dynamics (Golanov et al., 1994;

628    Karavaev et al., 2018; Zanatta et al., 2013) that influence the overall global dynamics of

629    neural excitability and connectivity (Fox et al., 2007, 2006; Fox and Raichle, 2007). Our

630    suggestion is that, insofar as the brain must model its own dynamic noise trajectories as a

631    function of active self-inference, non-neuronal sources of variability such as inscribed by

632    visceral rhythms must be incorporated within the brain's generative model of its own

633    percepts. Interoceptive fluctuations are thus an important influence over the precision of

634     exteroceptive sensory channels, and interoception is itself the means by which the brain

635     infers and controls its own pathway through these precision trajectories. The modelling

636     introduced here can thus be expanded beyond the cardiac domain to the more general

637     problem of modelling how spontaneous fluctuations in neurovisceral cycles (including

638     heart-rate variability) influence information processing and behavior.

639     What then, explains the lack of consistent results within the cardiac timing literature?

640     In contrast to the binary on/off hypotheses proposed by Lacey or Sokolov, our simulations

641     highlight the context-sensitive manner by which ascending visceral signals modulate the

642     precision of both interoceptive and exteroceptive inferences. For example, our simulation of

643     the startle response (illustrated in Fig. 3) clearly indicates that the functional impact of

644     cardio-ballistic responses is coupled to the agent's baseline prior expectations, as well as the

645     overall precision of active inference and policy selection. In this sense, whether a specific

646     cardiac response is likely to potentiate or inhibit a specific domain (e.g., fear) depends upon

647     the specific weighting of arousal policy priors, the precision of incoming exteroceptive and

648     interoceptive sensations, and the linkages thereof as determined by the task itself. In other

649     words, the specific balance of prior beliefs and sensory information, in a given cognitive or

650     affective domain, must be addressed before one can predict the exact directionality of an

651     interoceptive effect on perception, or *vice versa*. Here, we modelled the generation of arousal

652     policies as a function of hyper-parameters governing the preferred policy. In the future this

653     can be unpacked further by examining the divergence between prior and posterior beliefs

654     about these policies (e.g., through inferred epistemic value). Through Landauer's principle,

655     this divergence may be equated with the associated metabolic costs of computation and the

656     conceptual notion of interoceptive self-modelling (Kiverstein, 2018; Limanowski and

657     Blankenburg, 2013; Seth and Tsakiris, 2018).

658

659     *The computational neuroanatomy of interoceptive inference*

660     Having addressed the construct validity of our model, we now speculate as to some

661     likely neuronal substrates of the message passing implied by variational inference.

662     Interoceptive inference can be broken down into four core functional domains: basic

*Cardiac Active Inference*

663    sensory-motor control, conscious interoceptive (perceptual) awareness, metacognitive

664    monitoring, and hedonic (intrinsic) value. In our model, we focused primarily on the simplest

665    possible implementation of interoceptive inference, corresponding to the sensory-motor

666    domain (i.e., ascending and descending cardiac pathways) and their low-level interaction

667    with exteroceptive inference, via neuromodulatory gain control. Future work will benefit

668    from expanding upon our representation of uncertainty to include the computation of

669    epistemic and/or intrinsic value as proxies for these higher-order interoceptive systems

670    (Friston et al., 2017b; Parr and Friston, 2017).

671
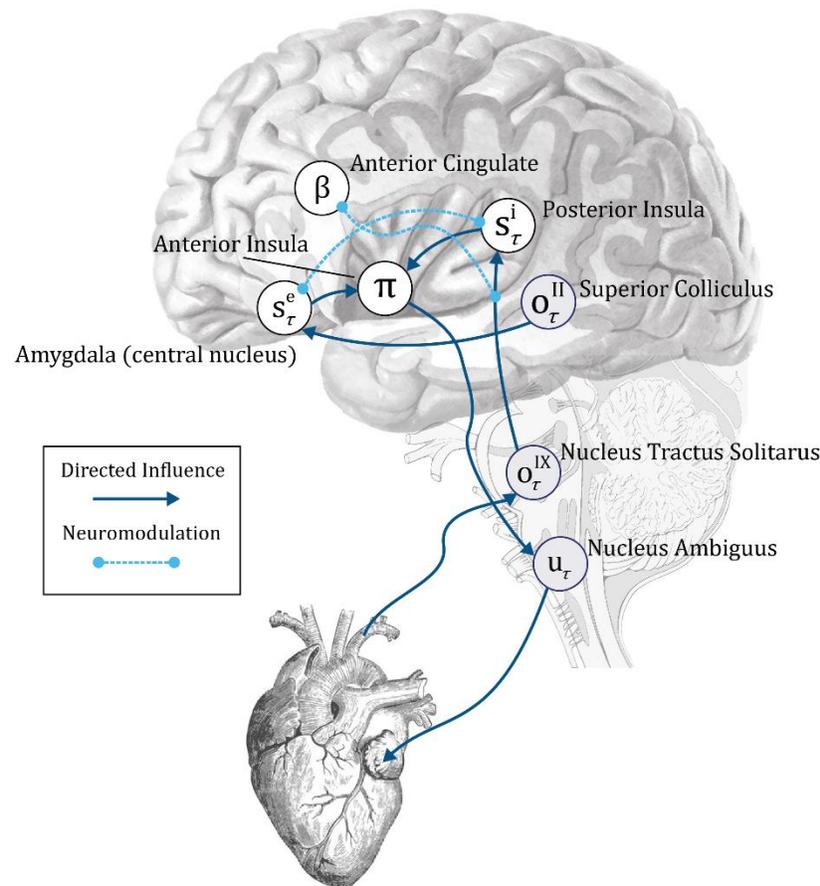


672

673    **Figure 6, computational neuroanatomy of interoception**. The schematic above shows the form of

674    the neuronal message passing implied by active inference for the generative model depicted in Figure 2. We

675    have related this to the anatomical networks that could implement these inferences. The sensory observations

676    in our simulations are visual and interoceptive (cardiac). These sensations are carried by cranial nerves II and

677    IX respectively. Cranial nerve II targets the superior colliculus in the midbrain. This structure sends short

678    latency visual data to the amygdala, which is well placed to make inferences about emotionally salient stimuli.

26

679  The amygdala additionally receives visual data from the ventral visual stream in the temporal lobe. Cranial

680  nerve IX carries information from the carotid sinus baroreceptors to the nucleus tractus solitarus in the

681  brainstem. This nucleus communicates with the posterior insula (via thalamic and PAG relays); the anterior

682  cingulate monitors and controls the precision of this ascending visceral information via neuromodulation,

683  possibly via feedback through noradrenergic pathways (not shown). The posterior insula and amygdala

684  interact with one another but also project to the anterior insula. This targets the nucleus ambiguus (via

685  brainstem relays such as the periaqueductal gray), which gives rise to the vagus (X) nerve. The vagus nerve

686  targets neurons in the cardiac plexus that project to both the sinoatrial node and the atrioventricular node of

687  the heart, slowing its rhythm. The nucleus tractus solitarus additionally participates in a reflex loop implicating

688  the sympathetic control of the cardiac cycle, but this is omitted for simplicity. The functional anatomy suggested

689  here implies the anterior insula might play a similar computational role in autonomic policy selection to the

690  basal ganglia in selection of policies involving the skeletal muscles (Friston et al., 2018). Note that inscribed

691  directed influences (blue arrows), are not assumed to be monosynaptic – for simplicity, many intermediary

692  relay nodes have been omitted.

693

694  Accordingly, in our sketch of the putative neuroanatomy underlying cardiac active

695  inference (Fig. 6), we focus primarily on the neuronal substrates that inscribe low-level

696  viscerosensory and visceromotor control, as well as some hierarchically superior regions

697  related to emotional salience and interoceptive awareness. For simplicity, our model depicts

698  only the minimal neuronal message passing scheme implied by our generative model; as

699  such, we have omitted many of the intermediary relay nodes; e.g., in the thalamus and ventral

700  visual stream. Afferent baroreceptor signals are transmitted along the ascending vagus to

701  the rostrum of the nucleus tractus solitarus (NTS, Mifflin and Felder, 1990; Miura and Reis,

702  1972). From here, ascending viscerosensory signal are projected via brainstem and midbrain

703  nuclei to the thalamus, somatosensory cortex, and posterior insula (Cechetto and Saper,

704  1987; Craig, 2002); ascending cardio-sensory outcomes are thus encoded in the NTS and

705  then passed to the posterior insular cortex (PIC) as inferred interoceptive states. The PIC has

706  a well-known role as primary viscerosensory cortex; electrical stimulation of this area elicits

707  phantom visceral sensations (e.g., pain, heart-rate acceleration) (Chouchou et al., 2019;

708  Oppenheimer et al., 1992) and bolus isoproterenol infusions increase the intensity of

709  cardiorespiratory sensations and concomitant PIC activations (Hassanpour et al., 2016;

710  Khalsa et al., 2009). In parallel, visual sensory outcomes are passed via the second cranial

711  nerve to the superior colliculus, where they inform exteroceptive inference in the amygdala,

712    which is well-situated to process salient emotional stimuli (Anderson and Phelps, 2001;
713    Liddell et al., 2005). These interoceptive and exteroceptive expectations then converge in the
714    anterior insular cortex (AIC), where they inform the selection of the appropriate autonomic
715    policy. Finally, the selected policy is passed down the hierarchy via descending pathways
716    (likely carried by von Economo neurons), to eventually engage the rostral nucleus ambiguus
717    and descending vagus, decelerating the heart-rate when the relaxed policy is selected.
718    Collectively, the scheme represents a multimodal reflex arc interlinking exteroceptive and
719    interoceptive domains to specific patterns of cardio-ballistic responses.

720          Within this scheme, we suggest that the rostral anterior cingulate (ACC) controls the
721    precision of ascending visceral outcomes and inferred interoceptive states via
722    neuromodulatory gain control (Fardo et al., 2017; Feldman and Friston, 2010). Further,
723    interoceptive and exteroceptive state precisions (in our scheme) interact indirectly through
724    global neuromodulatory influences, possibly through regulation of noradrenaline by the ACC
725    (via descending influence on the locus coeruleus). Neurobiologically and functionally
726    speaking, the AIC and ACC share similar profiles; both are densely populated with Von
727    Economo neurons (VENs), which are well-suited for the long-range modulation of neural
728    activity across the cortex (Allman et al., 2011), and also contain diverse populations of
729    noradrenergic, dopaminergic, and opioidergic neurons. Both regions further share an
730    integrative connectivity structure, with projections to both lower-level visceral-motor
731    brainstem nuclei and higher-order regions implicated in decision-making, metacognition,
732    and self-awareness, such as the ventromedial and dorsomedial prefrontal cortices (Allen et
733    al., 2017, 2016a; Fleming and Dolan, 2012; Menon and Uddin, 2010; Ullsperger et al., 2010).
734    However, the AIC is more densely interconnected with the PIC whereas the ACC is more
735    closely related to uncertainty and decision-making. On this basis, we propose that whereas
736    the AIC integrates the visceral and exteroceptive states required for the regulation of arousal
737    policies, the ACC is likely to regulate the gain or precision of these interactions[4].

---

[4] It is worth noting that this model may explain the widespread, seemingly unspecific activation profiles of these areas (Chang et al., 2013; Yarkoni et al., 2011), as the generative model specified here suggests both form part of an integrated hierarchical circuit by which interoceptive and exteroceptive states interact: e.g., either through the regulation of arousal policies or through the modulation of ascending viscerosensory precision.

738        What about metacognitive or reward-related interoceptive processes? Although here
739    we do not model these higher-order functions, the model can be expanded to include the
740    explicit representation of policy uncertainty and epistemic value as the mechanisms
741    underlying metacognitive self-inference; i.e., the integrative self-model that combines
742    exteroceptive and interoceptive predictions into a conscious schema (Allen and Tsakiris,
743    2019). In this case, we would expect that the VMPFC and DLPFC are likely to be engaged in
744    inferences about variables (e.g., those derived from expected free energy such as epistemic
745    and intrinsic value) that contextualize the inferences performed by the AIC and ACC over
746    longer timescales (Friston et al., 2015, 2017a).

747

*Limitations and Future Directions*

749    The model and simulations presented here represent a minimal proof-of-principle
750    demonstrating how cyclic interactions of interoceptive and exteroceptive perception arise
751    directly from the principles of active inference. Here, our primary goal was to move the
752    literature beyond purely conceptual analyses of 'interoceptive inference', to provide a formal
753    model sub-serving direct hypothesis testing. As such, we focus primarily on reproducing
754    commonly reported phenomena, rather than empirical cross-validation or biological
755    plausibility. While the model presented here does a reasonably good job of approximating
756    the cardiac cycle, it should be clear that much work remains to be done if the model is to be
757    used as a full generative model; e.g., of heart-brain interactions and/or physiological data
758    such as HRV. We therefore anticipate a variety of fruitful applications. For example, the
759    present MDP scheme could be expanded to include biologically realistic cardiac parameters,
760    or to include other visceral modalities such as gastric or respiratory fluctuations. Similarly,
761    the exteroceptive states modelled here could be adapted to a variety of experimental tasks
762    to capture embodied influences on, for example, active spatial navigation (Kaplan and
763    Friston, 2018; Lockmann et al., 2018; Lockmann and Tort, 2018), active reward learning
764    (FitzGerald et al., 2015; Marshall et al., 2019), interaction between the cardiac cycle and
765    ballistic saccades (Galvez-Pol et al., 2018; Mirza et al., 2016; Ohl et al., 2016), or
766    metacognitive self-inference (Allen et al., 2016b; Friston et al., 2017b; Hauser et al., 2017a).

767   These and other future directions will hopefully guide a newly embodied approach to

768   computational psychiatry, enabling the detailed phenotyping of clinical populations in terms

769   of aberrant interoceptive inference.

770

771

772

**Acknowledgements**

**Data and Code Availability**

The underlying MDP scheme here is available as part of the open-access distribution of SPM12. A demonstration of the scheme can be accessed by typing >>DEM into the Matlab command prompt and selecting the **Interoception** demo from the graphical user interface that appears. Further, all the code required to generate the simulations and figures herein can be found at the following github page: https://github.com/embodied-computation-group/cardiac-active-inference.

# References

Allen M, Fardo F, Dietz MJ, Hillebrandt H, Friston KJ, Rees G, Roepstorff A. 2016a. Anterior insula coordinates hierarchical processing of tactile mismatch responses. *NeuroImage* **127**:34–43. doi:10.1016/j.neuroimage.2015.11.030

Allen M, Frank D, Schwarzkopf DS, Fardo F, Winston JS, Hauser TU, Rees G. 2016b. Unexpected arousal modulates the influence of sensory noise on confidence. *eLife* **5**:e18103. doi:10.7554/eLife.18103

Allen M, Friston KJ. 2018. From cognitivism to autopoiesis: towards a computational framework for the embodied mind. *Synthese* **195**:2459–2482. doi:10.1007/s11229-016-1288-5

Allen M, Glen JC, Müllensiefen D, Schwarzkopf DS, Fardo F, Frank D, Callaghan MF, Rees G. 2017. Metacognitive ability correlates with hippocampal and prefrontal microstructure. *NeuroImage* **149**:415–423. doi:10.1016/j.neuroimage.2017.02.008

Allen M, Tsakiris M. 2019. The body as first prior: Interoceptive predictive processing and the primacyThe Interoceptive Mind: From Homeostasis to Awareness. Great Clarendon Street, Oxford, OX2 6DP: Oxford University Press. pp. 27–45.

Allman JM, Tetreault NA, Hakeem AY, Manaye KF, Semendeferi K, Erwin JM, Park S, Goubert V, Hof PR. 2011. The von Economo neurons in the frontoinsular and anterior cingulate cortex. *Ann N Y Acad Sci* **1225**:59–71. doi:10.1111/j.1749-6632.2011.06011.x

Anderson AK, Phelps EA. 2001. Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature* **411**:305–309. doi:10.1038/35077083

Apps MA, Tsakiris M. 2014. The free-energy self: a predictive coding account of self-recognition. *Neurosci Biobehav Rev* **41**:85–97.

Attias H. 2003. Planning by probabilistic inference.Proc. of the 9th Int. Workshop on Artifical Intelligence and Statistics. Presented at the AISTATS.

Azevedo RT, Garfinkel SN, Critchley HD, Tsakiris M. 2017. Cardiac afferent activity modulates the expression of racial stereotypes. *Nat Commun* **8**:13854. doi:10.1038/ncomms13854

Barto A, Mirolli M, Baldassarre G. 2013. Novelty or surprise? *Front Psychol* **4**:907.

Birn RM. 2012. The role of physiological noise in resting-state functional connectivity. *Neuroimage* **62**:864–870.

Boldt A, de Gardelle V, Yeung N. 2017. The impact of evidence reliability on sensitivity and bias in decision confidence. *J Exp Psychol Hum Percept Perform* **43**:1520–1531. doi:10.1037/xhp0000404

Bonvallet M, Bloch V. 1961. Bulbar control of cortical arousal. *Science* **133**:1133–1134.

Bonvallet M, Dell P, Hiebel G. 1954. Tonus sympathique et activité électrique corticale. *Electroencephalogr Clin Neurophysiol* **6**:119–144.

Botvinick M, Toussaint M. 2012. Planning as inference. *Trends Cogn Sci* **16**:485–488. doi:10.1016/j.tics.2012.08.006

Cechetto DF, Saper CB. 1987. Evidence for a viscerotopic sensory representation in the cortex and thalamus in the rat. *J Comp Neurol* **262**:27–45. doi:10.1002/cne.902620104

Chang LJ, Yarkoni T, Khaw MW, Sanfey AG. 2013. Decoding the Role of the Insula in Human Cognition: Functional Parcellation and Large-Scale Reverse Inference. *Cereb Cortex* **23**:739–749. doi:10.1093/cercor/bhs065

Chouchou F, Mauguière F, Vallayer O, Catenoix H, Isnard J, Montavont A, Jung J, Pichot V, Rheims S, Mazzola L. 2019. How the insula speaks to the heart: Cardiac responses to insular stimulation in humans. *Hum Brain Mapp*. doi:10.1002/hbm.24548

Cohen R, Lieb H, Rist F. 1980. Loudness judgments, evoked potentials, and reaction time to acoustic stimuli early and late in the cardiac cycle in chronic schizophrenics. *Psychiatry Res* **3**:23–29.

Craig AD. 2002. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat Rev Neurosci* **3**:655–666. doi:10.1038/nrn894

Delfini LF, Campos JJ. 1972. Signal Detection and the "Cardiac Arousal Cycle." *Psychophysiology* **9**:484–491. doi:10.1111/j.1469-8986.1972.tb01801.x

Denève S, Jardri R. 2016. Circular inference: mistaken belief, misplaced trust. *Curr Opin Behav Sci*, Computational modeling **11**:40–48. doi:10.1016/j.cobeha.2016.04.001

Eckberg Dwain L. 1997. Sympathovagal Balance. *Circulation* **96**:3224–3232. doi:10.1161/01.CIR.96.9.3224

Edwards L, Ring C, McIntyre D, Winer JB, Martin U. 2009. Sensory detection thresholds are modulated across the cardiac cycle: Evidence that cutaneous sensibility is greatest for systolic stimulation. *Psychophysiology* **46**:252–256. doi:10.1111/j.1469-8986.2008.00769.x

Elliott R. 1972. The significance of heart rate for behavior: A critique of Lacey's hypothesis.

Fardo F, Auksztulewicz R, Allen M, Dietz MJ, Roepstorff A, Friston KJ. 2017. Expectation violation and attention to pain jointly modulate neural gain in somatosensory cortex. *NeuroImage* **153**:109–121. doi:10.1016/j.neuroimage.2017.03.041

Feldman H, Friston K. 2010. Attention, Uncertainty, and Free-Energy. *Front Hum Neurosci* **4**:215. doi:10.3389/fnhum.2010.00215

FitzGerald THB, Dolan RJ, Friston K. 2015. Dopamine, reward learning, and active inference. *Front Comput Neurosci* **9**:136. doi:10.3389/fncom.2015.00136

Fleming SM, Dolan RJ. 2012. The neural basis of metacognitive ability. *Philos Trans R Soc B Biol Sci* **367**:1338–1349. doi:10.1098/rstb.2011.0417

Fox MD, Raichle ME. 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* **8**:700.

Fox MD, Snyder AZ, Vincent JL, Raichle ME. 2007. Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron* **56**:171–184.

Fox MD, Snyder AZ, Zacks JM, Raichle ME. 2006. Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nat Neurosci* **9**:23.

Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. 2015. Active inference and epistemic value. *Cogn Neurosci* **6**:187–214.

Friston KJ, FitzGerald T, Rigoli F, Schwartenbeck P, Pezzulo G. 2017a. Active inference: a process theory. *Neural Comput* **29**:1–49.

Friston KJ, Lin M, Frith CD, Pezzulo G, Hobson JA, Ondobaka S. 2017b. Active Inference, Curiosity and Insight. *Neural Comput* **29**:2633–2683. doi:10.1162/neco_a_00999

Friston KJ, Parr T, de Vries B. 2017c. The graphical brain: Belief propagation and active inference. *Netw Neurosci* **1**:381–414. doi:10.1162/NETN_a_00018

Friston KJ, Rosch R, Parr T, Price C, Bowman H. 2018. Deep temporal models and active inference. *Neurosci Biobehav Rev* **90**:486–501. doi:10.1016/j.neubiorev.2018.04.004

Gallagher S, Allen M. 2018. Active inference, enactivism and the hermeneutics of social cognition. *Synthese* **195**:2627–2648. doi:10.1007/s11229-016-1269-8

883  Galvez-Pol A, McConnell R, Kilner J. 2018. Active sampling during visual search is modulated
884      by the cardiac cycle. *bioRxiv* 405902.
885  Garfinkel SN, Critchley HD. 2016. Threat and the Body: How the Heart Supports Fear
886      Processing. *Trends Cogn Sci* **20**:34–46. doi:10.1016/j.tics.2015.10.005
887  Garfinkel SN, Minati L, Gray MA, Seth AK, Dolan RJ, Critchley HD. 2014. Fear from the Heart:
888      Sensitivity to Fear Stimuli Depends on Individual Heartbeats. *J Neurosci* **34**:6573–
889      6582. doi:10.1523/JNEUROSCI.3507-13.2014
890  Ghione S. 1996. Hypertension-associated hypalgesia: Evidence in experimental animals and
891      humans, pathophysiological mechanisms, and potential clinical consequences.
892      *Hypertension* **28**:494–504.
893  Golanov EV, Yamamoto S, Reis DJ. 1994. Spontaneous waves of cerebral blood flow
894      associated with a pattern of electrocortical activity. *Am J Physiol-Regul Integr Comp*
895      *Physiol* **266**:R204–R214. doi:10.1152/ajpregu.1994.266.1.R204
896  Graham FK, Clifton RK. 1966. Heart-rate change as a component of the orienting response.
897      *Psychol Bull* **65**:305–320. doi:10.1037/h0023258
898  Hassanpour MS, Yan L, Wang DJJ, Lapidus RC, Arevian AC, Simmons W. Kyle, Feusner Jamie
899      D., Khalsa Sahib S. 2016. How the heart speaks to the brain: neural activity during
900      cardiorespiratory interoceptive stimulation. *Philos Trans R Soc B Biol Sci*
901      **371**:20160017. doi:10.1098/rstb.2016.0017
902  Hauser TU, Allen M, Purg N, Moutoussis M, Rees G, Dolan RJ. 2017a. Noradrenaline blockade
903      specifically enhances metacognitive performance. *eLife* **6**. doi:10.7554/eLife.24901
904  Hauser TU, Allen M, Rees G, Dolan RJ. 2017b. Metacognitive impairments extend perceptual
905      decision making weaknesses in compulsivity. *Sci Rep* **7**:6614. doi:10.1038/s41598-
906      017-06116-z
907  Heathers JAJ. 2012. Sympathovagal balance from heart rate variability: an obituary. *Exp*
908      *Physiol* **97**:556–556. doi:10.1113/expphysiol.2011.063867
909  Herrero JL, Khuvis S, Yeagle E, Cerf M, Mehta AD. 2017. Breathing above the brainstem:
910      Volitional control and attentional modulation in humans. *J Neurophysiol*.
911  Itti L, Baldi P. 2009. Bayesian surprise attracts human attention. *Vision Res* **49**:1295–1306.
912  Kaplan R, Friston KJ. 2018. Planning and navigation as active inference. *Biol Cybern* **112**:323–
913      343. doi:10.1007/s00422-018-0753-2
914  Karavaev AS, Kiselev AR, Runnova AE, Zhuravlev MO, Borovkova EI, Prokhorov MD,
915      Ponomarenko VI, Pchelintseva SV, Efremova TY, Koronovskii AA, Hramov AE. 2018.
916      Synchronization of infra-slow oscillations of brain potentials with respiration. *Chaos*
917      *Interdiscip J Nonlinear Sci* **28**:081102. doi:10.1063/1.5046758
918  Khalsa SS, Rudrauf D, Sandesara C, Olshansky B, Tranel D. 2009. Bolus isoproterenol
919      infusions provide a reliable method for assessing interoceptive awareness. *Int J*
920      *Psychophysiol*, Central and peripheral nervous system interactions: From mind to
921      brain to body **72**:34–45. doi:10.1016/j.ijpsycho.2008.08.010
922  Kiverstein J. 2018. Free Energy and the Self: An Ecological–Enactive Interpretation. *Topoi*.
923      doi:10.1007/s11245-018-9561-5
924  Koch E. 1932. Die irradiation der pressoreceptorischen kreislaufreflexe. *J Mol Med* **11**:225–
925      227.
926  Kunzendorf S, Klotzsche F, Akbal M, Villringer A, Ohl S, Gaebler M. 2019. Active information
927      sampling varies across the cardiac cycle. *Psychophysiology* e13322.

928 Lacey BC, Lacey JI. 1978. Two-way communication between the heart and the brain:
929     Significance of time within the cardiac cycle. *Am Psychol* **33**:99.
930 Liddell BJ, Brown KJ, Kemp AH, Barton MJ, Das P, Peduto A, Gordon E, Williams LM. 2005. A
931     direct brainstem–amygdala–cortical 'alarm' system for subliminal signals of fear.
932     *NeuroImage* **24**:235–243. doi:10.1016/j.neuroimage.2004.08.016
933 Limanowski J, Blankenburg F. 2013. Minimal self-models and the free energy principle. *Front*
934     *Hum Neurosci* **7**. doi:10.3389/fnhum.2013.00547
935 Lockmann ALV, Laplagne DA, Tort ABL. 2018. Olfactory bulb drives respiration-coupled beta
936     oscillations in the rat hippocampus. *Eur J Neurosci* **48**:2663–2673.
937     doi:10.1111/ejn.13665
938 Lockmann ALV, Tort ABL. 2018. Nasal respiration entrains delta-frequency oscillations in
939     the prefrontal cortex and hippocampus of rodents. *Brain Struct Funct* **223**:1–3.
940     doi:10.1007/s00429-017-1573-1
941 Malliani A, Pagani M, Lombardi F, Cerutti S. 1991. Cardiovascular neural regulation explored
942     in the frequency domain. *Circulation* **84**:482–492. doi:10.1161/01.CIR.84.2.482
943 Marshall AC, Gentsch A, Blum A-L, Broering C, Schütz-Bosbach S. 2019. I feel what I do:
944     Relating interoceptive processes and reward-related behavior. *NeuroImage*
945     **191**:315–324. doi:10.1016/j.neuroimage.2019.02.032
946 Menon V, Uddin LQ. 2010. Saliency, switching, attention and control: a network model of
947     insula function. *Brain Struct Funct* **214**:655–667. doi:10.1007/s00429-010-0262-0
948 Mifflin SW, Felder RB. 1990. Synaptic mechanisms regulating cardiovascular afferent inputs
949     to solitary tract nucleus. *Am J Physiol-Heart Circ Physiol* **259**:H653–H661.
950     doi:10.1152/ajpheart.1990.259.3.H653
951 Mirza MB, Adams RA, Mathys CD, Friston KJ. 2016. Scene Construction, Visual Foraging, and
952     Active Inference. *Front Comput Neurosci* **10**:56. doi:10.3389/fncom.2016.00056
953 Miura M, Reis DJ. 1972. The role of the solitary and paramedian reticular nuclei in mediating
954     cardiovascular reflex responses from carotid baro- and chemoreceptors. *J Physiol*
955     **223**:525–548. doi:10.1113/jphysiol.1972.sp009861
956 Ohl S, Wohltat C, Kliegl R, Pollatos O, Engbert R. 2016. Microsaccades are coupled to
957     heartbeat. *J Neurosci* **36**:1237–1241.
958 Oppenheimer SM, Gelb A, Girvin JP, Hachinski VC. 1992. Cardiovascular effects of human
959     insular cortex stimulation. *Neurology* **42**:1727. doi:10.1212/WNL.42.9.1727
960 Oudeyer P-Y, Kaplan F. 2009. What is intrinsic motivation? A typology of computational
961     approaches. *Front Neurorobotics* **1**:6.
962 Owens AP, Allen M, Ondobaka S, Friston KJ. 2018. Interoceptive inference: From
963     computational neuroscience to clinic. *Neurosci Biobehav Rev* **90**:174–183.
964     doi:10.1016/j.neubiorev.2018.04.017
965 Park H-D, Correia S, Ducorps A, Tallon-Baudry C. 2014. Spontaneous fluctuations in neural
966     responses to heartbeats predict visual detection. *Nat Neurosci* **17**:612–618.
967     doi:10.1038/nn.3671
968 Park H-D, Tallon-Baudry C. 2014. The neural subjective frame: from bodily signals to
969     perceptual consciousness. *Philos Trans R Soc B Biol Sci* **369**:20130208.
970 Parr T, Friston KJ. 2018. The Computational Anatomy of Visual Neglect. *Cereb Cortex* **28**:777–
971     790. doi:10.1093/cercor/bhx316
972 Parr T, Friston KJ. 2017. Uncertainty, epistemics and active inference. *J R Soc Interface*
973     **14**:20170376. doi:10.1098/rsif.2017.0376

Parr T, Markovic D, Kiebel SJ, Friston KJ. 2019. Neuronal message passing using Mean-field, Bethe, and Marginal approximations. *Sci Rep* **9**:1889. doi:10.1038/s41598-018-38246-3

Peters A, McEwen BS, Friston K. 2017. Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. *Prog Neurobiol* **156**:164–188. doi:10.1016/j.pneurobio.2017.05.004

Petzschner FH, Weber LAE, Gard T, Stephan KE. 2017. Computational Psychosomatics and Computational Psychiatry: Toward a Joint Framework for Differential Diagnosis. *Biol Psychiatry*, Computational Psychiatry **82**:421–430. doi:10.1016/j.biopsych.2017.05.012

Powers AR, Mathys C, Corlett PR. 2017. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science* **357**:596–600. doi:10.1126/science.aan3458

Salomon R, Ronchi R, Dönz J, Bello-Ruiz J, Herbelin B, Martet R, Faivre N, Schaller K, Blanke O. 2016. The Insula Mediates Access to Awareness of Visual Stimuli Presented Synchronously to the Heartbeat. *J Neurosci* **36**:5115–5127. doi:10.1523/JNEUROSCI.4262-15.2016

Sandman CA, McCanne TR, Kaiser DN, Diamond B. 1977. Heart rate and cardiac phase influences on visual perception. *J Comp Physiol Psychol* **91**:189.

Saxon SA. 1970. Detection of near threshold signals during four phases of cardiac cycle. *Ala J Med Sci* **7**:427.

Schmidhuber J. 2010. Formal theory of creativity, fun, and intrinsic motivation. *IEEE Trans Auton Ment Dev* **2**:230–247.

Seth AK. 2013. Interoceptive inference, emotion, and the embodied self. *Trends Cogn Sci* **17**:565–573. doi:10.1016/j.tics.2013.09.007

Seth AK, Friston KJ. 2016. Active interoceptive inference and the emotional brain. *Philos Trans R Soc B Biol Sci* **371**:20160007. doi:10.1098/rstb.2016.0007

Seth AK, Tsakiris M. 2018. Being a Beast Machine: The Somatic Basis of Selfhood. *Trends Cogn Sci* **22**:969–981. doi:10.1016/j.tics.2018.08.008

Sokolov EN. 1963. Higher nervous functions: The orienting reflex. *Annu Rev Physiol* **25**:545–580.

Spence ML, Dux PE, Arnold DH. 2016. Computations underlying confidence in visual perception. *J Exp Psychol Hum Percept Perform* **42**:671–682. doi:10.1037/xhp0000179

Strigo A, Craig AD. 2016. Interoception, homeostatic emotions and sympathovagal balance. *Philos Trans R Soc B Biol Sci* **371**:20160010. doi:10.1098/rstb.2016.0010

Tort ABL, Brankačk J, Draguhn A. 2018a. Respiration-Entrained Brain Rhythms Are Global but Often Overlooked. *Trends Neurosci* **41**:186–197. doi:10.1016/j.tins.2018.01.007

Tort ABL, Ponsel S, Jessberger J, Yanovsky Y, Brankačk J, Draguhn A. 2018b. Parallel detection of theta and respiration-coupled oscillations throughout the mouse brain. *Sci Rep* **8**:6432–6432. doi:10.1038/s41598-018-24629-z

Ullsperger M, Harsay HA, Wessel JR, Ridderinkhof KR. 2010. Conscious perception of errors and its relation to the anterior insula. *Brain Struct Funct* **214**:629–643. doi:10.1007/s00429-010-0261-1

1018 Varga S, Heck DH. 2017. Rhythms of the body, rhythms of the brain: Respiration, neural
1019        oscillations, and embodied cognition. *Conscious Cogn* **56**:77–90.
1020        doi:10.1016/j.concog.2017.09.008
1021 Velden M, Juris M. 1975. Perceptual Performance as a Function of Intra-Cycle Cardiac
1022        Activity. *Psychophysiology* **12**:685–692. doi:10.1111/j.1469-8986.1975.tb00075.x
1023 Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale automated
1024        synthesis of human functional neuroimaging data. *Nat Methods* **8**:665–670.
1025        doi:10.1038/nmeth.1635
1026 Zanatta P, Toffolo GM, Sartori E, Bet A, Baldanzi F, Agarwal N, Golanov E. 2013. The human
1027        brain pacemaker: Synchronized infra-slow neurovascular coupling in patients
1028        undergoing non-pulsatile cardiopulmonary bypass. *NeuroImage* **72**:10–19.
1029        doi:10.1016/j.neuroimage.2013.01.033
1030 Zelano C, Jiang H, Zhou G, Arora N, Schuele S, Rosenow J, Gottfried JA. 2016. Nasal Respiration
1031        Entrains Human Limbic Oscillations and Modulates Cognitive Function. *J Neurosci*
1032        **36**:12448–12467. doi:10.1523/JNEUROSCI.2586-16.2016
1033
1034