



Interoceptive inference: From computational neuroscience to clinic

Andrew P. Owens^{a,b,c,*}, Micah Allen^{d,e}, Sasha Ondobaka^{e,f}, Karl J. Friston^d

^a Lab of Action & Body, Department of Psychology, Royal Holloway, University of London, Egham, Surrey, UK

^b Department of Brain Repair and Rehabilitation, Institute of Neurology, University College London, UK

^c Autonomic Unit, National Hospital Neurology and Neurosurgery, UCL NHS Trust, London, UK

^d Wellcome Trust Centre for Neuroimaging, University College London, UK

^e Institute of Cognitive Neuroscience, University College London, UK

^f Sobell Department for Motor Neuroscience and Movement Disorders, University College London, UK

ARTICLE INFO

Keywords:

Active inference
Autonomic nervous system
Free-energy principle
Homeostasis
Interoception
Interoceptive inference
Predictive coding

ABSTRACT

The central and autonomic nervous systems can be defined by their anatomical, functional and neurochemical characteristics, but neither functions in isolation. For example, fundamental components of autonomically mediated homeostatic processes are afferent interoceptive signals reporting the internal state of the body and efferent signals acting on interoceptive feedback assimilated by the brain. Recent predictive coding (interoceptive inference) models formulate interoception in terms of embodied predictive processes that support emotion and selfhood. We propose interoception may serve as a way to investigate holistic nervous system function and dysfunction in disorders of brain, body and behaviour. We appeal to predictive coding and (active) interoceptive inference, to describe the homeostatic functions of the central and autonomic nervous systems. We do so by (i) reviewing the active inference formulation of interoceptive and autonomic function, (ii) survey clinical applications of this formulation and (iii) describe how it offers an integrative approach to human physiology; particularly, interactions between the central and peripheral nervous systems in health and disease.

1. Introduction

‘Interoception’ refers to afferent sensory information arising from the sensation, perception, and awareness of afferent feedback from the viscera that underwrites homeostatic functioning (Craig, 2002). The control of interoceptive stability or homeostasis (i.e., autonomic nervous system regulation) can be mapped onto a hierarchical organisation; ranging from basic physiological reflexes to global cortical networks that integrate the function of the central and autonomic nervous systems (Owens et al., 2017a). Fundamental components of these homeostatic processes are afferent interoceptive signals reporting the internal state of the body and efferent signals acting on interoceptive feedback (Barrett and Simmons, 2015; Ondobaka et al., 2015a; Quattrocki and Friston, 2014; Park et al., 2014), in the form of homeostatic reflexes that are informed by somatic states represented in the central nervous system. Co-ordinated central and peripheral nervous system function is required, even at lower tiers in the hierarchy, where structures such as the spinal cord, brainstem and hypothalamus mediate autonomic outflows and descending cortical inhibition (Calejesan et al., 2000; Benarroch, 1993). For example, the periaqueductal gray (PAG), which regulates input/output of nociceptive and visceral signals, is also innervated by descending anterior cingulate

cortex (ACC) projections, which can boost or inhibit pain responsivity, selectively (Calejesan et al., 2000). Moreover, chemoreceptors in the brain stem monitor arterial carbon dioxide, oxygen and hydrogen ion levels to regulate carbon dioxide, oxygen and pH perfusion via sympathetic and phrenic efferents. More generally, hypothalamic, pontine and medullary sympathetic and parasympathetic nuclei interact with homeostatic representations to generate effector-organ specific autonomic responses (Saper, 2002). In the cardiovascular domain, heart rate changes are related to activity in the amygdala and dorsal anterior cingulate cortex (dACC) (Janig and Habler, 2003) and during stress, amygdala activity predicts systolic contractility (Dalton et al., 2005). The amygdala, ACC and other limbic structures supply descending inputs to the hypothalamus and brainstem for emotion-related autonomic responses (Saper, 2002).

1.1. The functional anatomy of interoception

As key players in the functional anatomy of interoception, the ACC and insula cortex are important for the processing of interoceptive feedback and mediating autonomic responses to interoceptive information (Medford and Critchley, 2010; Damasio and Carvalho, 2013). dACC (Critchley et al., 2003) and insula cortex (Critchley et al., 2000a;

* Corresponding author at: Lab of Action & Body, Department of Psychology, Royal Holloway, University of London, Egham, Surrey, TW20 0EXU, UK.
E-mail address: andrew.owens@rhul.ac.uk (A.P. Owens).

Critchley et al., 2000b) activity reflects engagement of sympathetic nervous system activity coupled to mental and physical behaviours. The anterior and posterior insula show increased neuronal activity during respiration, isometric exercise, cold pressor and Valsalva manoeuvres (King et al., 1999; Harper et al., 2000). Increases in blood pressure positively correlate with right dACC activity (Critchley et al., 2000a), supporting findings that sympathetic responses are lateralized to the right hemisphere (Oppenheimer et al., 1992), whereas the left insular cortex is involved in parasympathetic nervous system cardiovascular regulation, as evidenced by acute left insular stroke disrupting the correlation between heart rate and blood pressure (Oppenheimer et al., 1996).

The insula has a posterior-to-anterior gradient, with initial sensory afferent information received by the posterior insula, which is then passed to the anterior insula cortex (AIC) – especially the right – where it is integrated with cognitive-affective biases and autobiographical information. This unique integrative structure has led to a variety of models relating to the function of the region, ranging from general theories of consciousness and affect to a putative role as a primary viscerosensory region (Klein et al., 2013). Accordingly, the AIC modulates homeostatic autonomic and interoceptive function via connections to allostatic centres (Flynn, 1999). Reduced baroreceptor tone is associated with ACC, amygdala and AIC function, whereas initiation of baroreflexes increases activity in lateral prefrontal cortex (IPFC) and posterior insula (Kimmerly et al., 2005). The mid and posterior insula are associated with somatomotor function and representations (Damasio et al., 2000) and the AIC and mid insula cortices, ACC and somatomotor cortex are functionally associated with shifting one's attention to interoceptive signals (Critchley et al., 2004). Bilateral insula cortices are activated during oesophageal stimulation (Binkofski et al., 1998) but as stimulation increases to the point of becoming painful, the right AIC is recruited (Aziz et al., 1997), illustrating how increasing interoceptive feedback will ascend the interoceptive hierarchy from bilateral insula to right AIC, as initial reporting of somatic sensory feedback escalates to a violation of homeostasis then to nociception; engaging conscious awareness. More generally, the insula is implicated in the integration of both interoceptive and exteroceptive inputs, and has been proposed to act as a core comparator underlying the generation of a multisensory embodied self (Allen et al., 2016a; Allen and Friston, 2016), which also regulates interactions between the cognitive and affective aspects of pain (Singer et al., 2009; Wiech et al., 2010; Fardo et al., 2015).

With respect to descending neural pathways, central efferent signals can drive allostatic changes in autonomic and behavioural function. During rest (Nakamura et al., 2008) and exercise (Tattersson et al., 2000; Tucker et al., 2006), perceived changes in skin temperature and thermal discomfort typically induce behavioural modifications before the recruitment of endocrine or autonomic thermostatic mechanisms (Schlader et al., 2009; Mundel et al., 2007). Behaviour-dependent increases in blood pressure are enabled and moderated by the baroreflex (Dampney et al., 2013; Dampney et al., 2002) and baroreflex dysfunction causes loss of consciousness due to cerebral hypoperfusion. The baroreflex arc ensures cerebral perfusion by mechanoreceptors in the carotid arteries and aortic arch detecting changes in arterial pressure and constantly feeding back this interoceptive information to the nucleus of the solitary tract (NTS), which synapses with the rostral ventrolateral medulla to set efferent pressor tone. During emotional or cognitive stress, the baroreflex feedback loop is disrupted by top-down cortical influences, increasing heart rate and blood pressure during steady-state physiological demands. Specifically, the aberrant cardiovascular up-regulation in the absence of allostatic demand results from suppression of low-order baroreceptor brainstem signalling by the solitary nucleus of the medulla, hippocampus, hypothalamic nuclei and prefrontal cortex (PFC) (Skinner, 1988). In summary, although the central and autonomic nervous systems are defined by unique anatomical, functional and neurochemical characteristics, they also interact

in a variety of ways to maintain homeostasis. Interoceptive signalling and control spans and integrates central and peripheral homeostatic processes, as well as influencing emotional and cognitive functions (Damasio, 1999; Gray et al., 2012; Lange and James, 1922).

In the following, we propose that interoception may provide a unique window into holistic human nervous system function and dysfunction in disorders of brain, body and behaviour. Due to the scope of this proposition, we offer a formal framework – grounded in interoceptive inference – that offers a methodological foundation for generating empirical predictions. To this end, we first formulate homeostasis in terms of interoceptive inference; *via* symbiotic interoceptive and autonomic nervous system function, before describing the clinical application of this approach. We then illustrate how this formulation can offer an overarching approach to human physiology, particularly autonomically mediated systems. Finally, we will review our initial empirical findings and their relationship to interoceptive inference.

1.2. Interoceptive predictive coding – neural correlates for conscious and unconscious processes

Discrepancies between predicted and experienced interoceptive signals have been proposed as a potential cause for anxiety (Paulus and Stein, 2006). In predictive coding terms, discrepancies between 'top-down' predictions generated by the brain and incoming sensory signals from the periphery are compared to produce a 'prediction error'. Subsequent minimisation of this prediction error corresponds to a Bayes optimal estimation of how sensory signals were caused; this can be seen easily by noting that if descending predictions match sensations exactly, the predictions must have been generated by representations of the world (*i.e.* expectations) that are, in some sense, veridical. This can be formalised in terms of Bayesian inference, where the evaluation of an expectation about the world is based on prior beliefs and the likelihood of observed data.

The application of predictive coding to perceptual inference involves minimisation of unpredicted or surprising sensory signals (prediction errors) within the cortical hierarchy by the generation of top-down predictions (Fig. 1). In this setting, the prediction errors at the sensory level play the role of a likelihood (*i.e.*, reporting how unlikely the sensations were given expectations about their causes), while prediction errors at higher levels play the role of empirical priors (*i.e.*, how unlikely expectations at one level are, given expectations of the level above). It is fairly easy to show that minimising prediction errors at each and every level of the hierarchy produces a set of expectations that constitute a Bayes optimal representation of how sensations are generated (Rao and Ballard, 1999; Friston, 2008, 2010). In brief, the minimisation of prediction errors involves reciprocal exchange of signals between hierarchical levels: prediction errors ascend the hierarchy to revise expectations, which generate descending predictions that resolve or suppress prediction errors at the level below.

In biologically plausible versions of the scheme (Friston, 2008, 2010; Shipp, 2016), prediction errors are thought to be encoded by the activity of superficial pyramidal neurons, which compare expectations with predictions descending from deep-layer pyramidal neurons in higher hierarchical levels. The prediction error is then projected (*via* intrinsic or interlaminar connections) to deep pyramidal cells encoding expectations in the higher cortical level, enabling a more accurate prediction to be reciprocated. This recurrent message passing allows prediction units to produce a more accurate prediction and effectively silence prediction error.

A prediction error's strength or influence on expectations or representations as higher levels depends on its 'precision' or reliability (Fig. 1). If a prediction error is less reliable, such as vision on a foggy day, more precision or weight will be afforded to prior expectations or beliefs about the environment. This ensures Bayes optimal perception, meaning that precision determines the influence of prediction error on subsequent hierarchical cortical evidence (*i.e.*, prediction error)

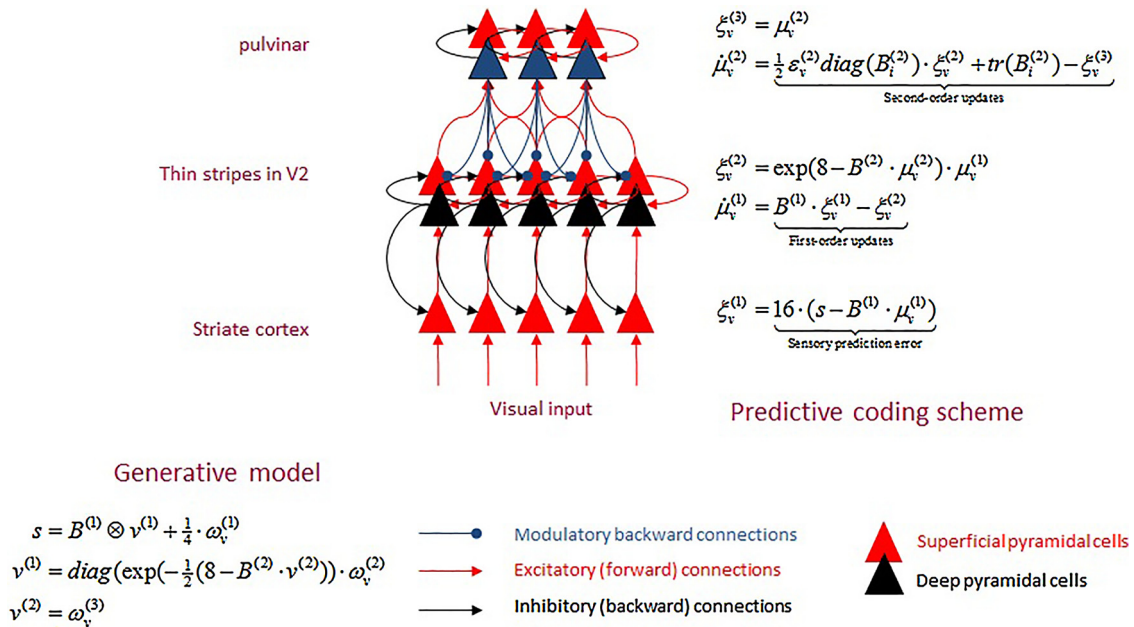


Fig. 1. This schematic illustrates the message passing implicit in predictive coding based on the generative model described (mathematically) on the lower left. Sensory input is conveyed to visual cortex via ascending prediction errors from the lateral geniculate nucleus. Posterior expectations, encoded by the activity of deep pyramidal cells in primary visual cortex, are driven by ascending prediction errors while, at the same time, they are subject to lateral interactions – with second level prediction errors – that mediate (empirical) priors. These constraints are modulated by top-down predictions of their precision (blue arrows). These predictions are based upon expectations about precision in the highest level that are effectively driven by the variance or power of prediction errors at the lower level. Heuristically, expectations about precision release posterior expectations from constraints in the vicinity of an inferred object and allow them to respond more sensitively to ascending geniculate input (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

accumulation. This hierarchical form of estimation for inference necessitates a generative model, in which the expected cause of representations at one level of the hierarchy become priors for expectations in the subordinate level. The term ‘generative model’ is used because the model generates the predictions of subordinate causes and ultimately sensations *per se*. When a generative (*i.e.* internal or forward) model is converted given the data at hand, sensations are explained in terms of the most likely hierarchical causes. These expected causes are constantly updated as new data are successively sampled to provide a biologically plausible form of evidence accumulation for data assimilation (Kanai et al., 2015). The inversion of a generative model means inferring the causes (*i.e.*, hidden states of the world) from consequences (*i.e.*, sensory samples from the world). This inverts the mapping prescribed by the generative model that generates consequences (sensations) from causes (hidden or latent states) (Friston, 2008). This mapping is also known as a forward model.

The primate brain is hierarchically structured (Felleman and Van Essen, 1991), which suggests the generative model used by the brain must also be hierarchical. This hierarchical architecture allows for the reciprocal message passing of predictions and prediction errors among hierarchical levels described above. Predictive coding models – derived from the ‘the free-energy principle’ (FEP) (Friston, 2009; Friston, 2010) – assume the brain endeavours to minimise precision weighted prediction errors throughout and implicitly maximise the evidence for its generative model.¹ This is known as self-evidencing (Hohwy, 2016), which can be regarded as a generalisation of homeostasis to every sensory modality predicted by the brain. Crucially, the FEP posits a defining role for homeostatic and allostatic processes in the functioning of the nervous system by casting the homeostatic imperative to stay alive as an innate and very precise prior over physiological states (Singer et al., 2009).

¹ Free energy can be regarded as the total amount of (precision weighted) prediction error summed over all levels of a hierarchical model.

1.3. Active inference under the free-energy principle

Under the free-energy principle, ‘active inference’ refers to the bilateral reduction of free-energy when: (i) prediction errors ascend the cortical hierarchy to change predictions, or (ii) prediction errors descend to the periphery to engage motor (or autonomic) reflexes, which change sensations (Ondobaka et al., 2015b). In the sensory system, prediction errors can only be modified by changing predictions, whereas proprioceptive and interoceptive prediction errors can also be modified by engaging reflexes to alter the sensory signal at its point of origin. In short, the prediction error can be reduced by changing the prediction (*i.e.*, perception) or by changing the sensations being predicted (*i.e.*, action). Movements can be initiated by predictions of the sensory consequences of action because the motor system automatically moves the sense organs to meet proprioceptive predictions, thereby shifting the imperative for action from what the individual wants to achieve with the action to what he/she wants to experience (Friston, 2010). An intuitive example of this is the common pain reflex; if a sufficiently precise and unexpected stimulus is received (*e.g.*, placing one’s hand on a hot stove), an optimal response would be to immediately alter sensations (*e.g.*, by withdrawing one’s hand); rather than to update one’s beliefs about the stimulus. This would be the homologue of allostasis that calls on deep or hierarchical inference – driven by ascending prediction errors – to restore homeostasis. One can see that the two ways of minimising prediction error depend sensitively on the precision afforded to *ascending* prediction errors, in relation to *descending* prediction errors. In this setting, ascending prediction errors ascend from the periphery to update central representations or expectations, while descending prediction errors are directed back to the periphery and effector systems. In the motor system, descending prediction errors would correspond to efferent (alpha motor neuron) discharge targeting the neuromuscular junction and eliciting a reflex. In short, the balance of precision or gain afforded prediction errors determine whether they are directed centrally to revise beliefs and elicit allostatic responses – or directed peripherally to elicit reflexes.

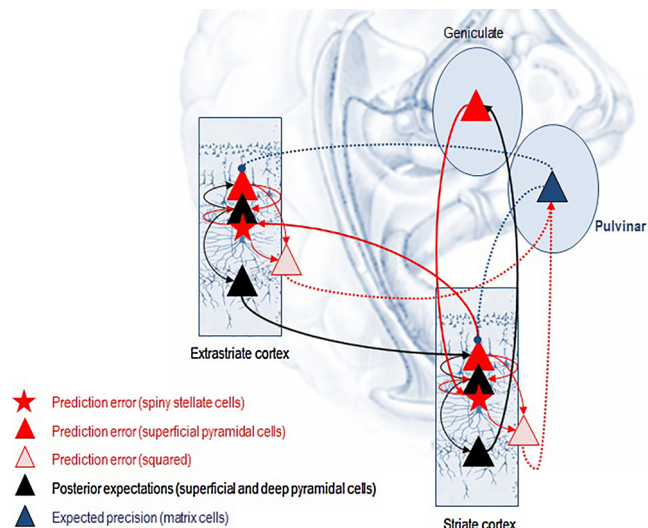


Fig. 2. This schematic details putative laminar-specific connections that are consistent with the precision-based predictive coding scheme in the main text. This architecture conforms roughly to the known neuroanatomy and physiology of canonical microcircuits and laminar specificity of extrinsic connections. The key aspect of this figure is the inclusion of deep pyramidal cells encoding the amplitude of prediction error (squared) that inform posterior expectations about precision in the (matrix cells) of the pulvinar. These cells reciprocate descending projections to modulate the gain of superficial pyramidal cells in cortex. Forward connections are in red and descending (backward) connections are in black. First-order streams are shown as full lines and second-order (precision related) streams are shown as broken lines (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

This functional architecture means that one can either ignore ascending prediction errors – via sensory attenuation – to engage reflexes (c.f., homeostasis) or allow ascending prediction errors to engage adaptive behaviour (c.f., allostasis). For example, if I inadvertently elicited a reflexive withdrawal (e.g., by grasping a hot spoon). This

withdrawal would be driven by descending prediction errors that do not update central beliefs. In other words, I will be aware that I have dropped the spoon" but may not "feel the pain". Conversely, if nociceptive input persists (e.g., while carrying a hot cup of tea), the initial attenuation of ascending prediction errors will abate, causing me to revise my beliefs and attend to a careful and allostatic response (e.g., replacing it quickly and safely on some suitable surface). In what follows, we look more closely at the crucial role of precision in mediating between these two sorts of responses.

1.4. Precision and gain control

The precision of ascending prediction errors determines the balance between priors and sensory signals to govern the influence of sensory evidence and prior beliefs (Fig. 2). In this setting, attention is intimately related to precision; in that attention is thought to increase the precision of prediction errors so that they have a greater influence on perception. Conversely, sensory attenuation is thought to reduce the precision of ascending prediction errors to enable motor reflexes to be driven by descending prediction errors. In other words, sensory attenuation involves ignoring the consequences of action, so that precise predictions (*i.e.* intentions) are realised by directing (descending) prediction errors to effector organs (Brown et al., 2013; Hughes et al., 2013; Wiese, 2017); *i.e.*, changing the state of the world (as opposed to revising our beliefs about the world). This sensory attenuation is generally thought to be pre-emptive and transient (Blakemore et al., 1999). Perhaps the clearest example is saccadic suppression (Wurtz, 2008); namely, the attenuation of ascending visual prediction errors during saccadic eye movements that last for 100 ms or so. The attenuation is then suspended so that we can attend to the sensory information (*i.e.*, ascending prediction errors) we have chosen to sample. In this scenario, by allowing predictions to be fulfilled via spinal reflex arcs, sensory attenuation allows movement to occur. On this view, sensory attenuation is a necessary precondition for – and part of – an intended movement (Parees et al., 2014) – in contrast to formulations in which sensory attenuation is crucial for a *post hoc* labelling of movements as self-generated (Blakemore et al., 1999). A key neurobiological issue here is

Precision and gain control

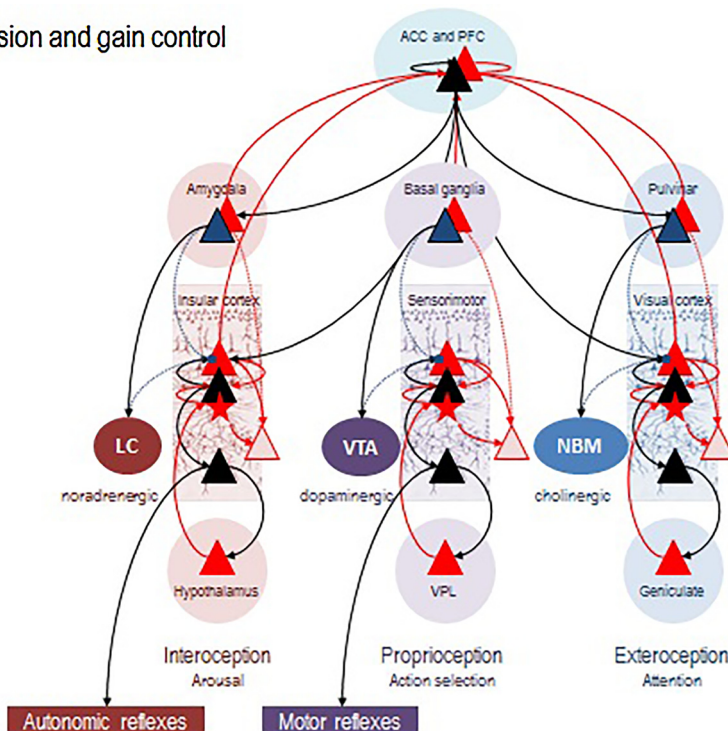


Fig. 3. This schematic extends the pulvinar example to provide a rough sketch of equivalent precision or gain control in interoceptive and proprioceptive systems. The architecture and anatomical designations should not be taken too seriously. However, there may be homologous architectures for exteroceptive, proprioception and interoception. Here, we have indicated this by assigning the pulvinar, basal ganglia and amygdala a common role; namely to provide precision control or contextual guidance to interoceptive (insular), proprioceptive (sensorimotor) and exteroceptive (visual) cortex respectively. In addition, each of these systems has been associated with a specific neuromodulator; namely noradrenaline, dopamine and acetylcholine in the ensuing regulation of autonomic arousal, action selection and attentional selection, respectively. Crucially, in a hierarchical setting, all these domain specific systems are integrated at the levels of the hierarchy (here attributed to the anterior cingulate and prefrontal cortex). Note that the recurrent or reciprocal message passing means that changes in the precision or postsynaptic gain in one (e.g., interoceptive) system, will necessarily effect processing in the others (e.g., exteroceptive). This is a necessary consequence of Bayes optimal inference in the sorts of hierarchical models. Note that the only way that this inference can act upon the world is through autonomic or motor reflexes. This means that exteroceptive processing has to be hierarchically integrated with proprioceptive and interoceptive inference – so that it can contextualise behaviour LC, locus coeruleus. VTA, ventral tegmental area. NBM, Nucleus Basalis of Meynert. VPL, ventral posterolateral thalamus. ACC, anterior cingulate cortex. PFC, prefrontal cortex.

that precision can be associated with the excitability or postsynaptic gain of units encoding prediction error (Moran et al., 2013; Aukstulewicz and Friston, 2015; Bauer et al., 2014; Brown and Friston, 2012; FitzGerald et al., 2015).

2. Interoceptive (active) inference in theory and practice

We now briefly describe how active inference may be transferred from the proprioceptive to the interoceptive domain as interoceptive inference. This will help elucidate how interoceptive inputs drive the autonomic nervous system to mediate homeostasis (Fig. 3). We begin with psychophysiological aspects, such as classical conditioning, the placebo effect and substance abuse and relapse, as well as affective disorders and psychosomatic illness. We then describe how interoceptive inference can offer an overarching methodology to study human physiology; using bladder function and thermoregulation as examples – in addition to the previously described cardiovascular reflex arcs. We conclude with a brief summary of our studies of interoceptive inference.

The free-energy principle has been applied to proprioceptive and exteroceptive sensory domains to elucidate the neurobiology of perception, motor control and attention (Aukstulewicz and Friston, 2015; Vossel et al., 2015), with applications to autism spectrum disorder (Gu et al., 2015a) and schizophrenia (Fogelson et al., 2014; Bastos et al., 2015). However, its potential role in interoception has only recently been considered (Barrett and Simmons, 2015; Ondobaka et al., 2015a; Quattrocki and Friston, 2014; Seth et al., 2011; Seth and Critchley, 2013). It has been suggested that Pavlovian classical conditioning can be viewed an elementary form of interoceptive inference (Pezzulo et al., 2015). Pavlov demonstrated not only that an unconditioned interoceptive prediction error (food) induces homeostatic autonomic responses (salivation) but that through the encoding of an exteroceptive signal (a bell), the same autonomic reflex can be induced by top-down predictions (Pavlov, 1927). Recently we have found empirical support for interoceptive inference by demonstrating that the orienting response, which was first described by Pavlov, is exaggerated during combined emotional aversion and interoceptive threat (and therefore, increased interoceptive prediction error). This evidence was provided by studies of dysautonomic symptom provocation in patients with postural tachycardia syndrome and vasovagal syncope – two forms of dysautonomia defined by baroreflex dysfunction (Owens et al., 2015a; Owens et al., 2018a). These findings provide insights into how interoceptive inference can prescribe autonomic reflexes and the destructive effect of dysautonomia on homeostasis – due to the breakdown of autonomic reflex arcs.

Pavlov also foreshadowed the role of predictive coding when noting that previously neutral stimuli conditioned the effects of apomorphine and morphine (Pavlov, 1960). The involvement of the opiate system in the placebo effect (Levine et al., 1978) further suggests that interoceptive inference can explain how inert stimuli can induce physiological responses *via* the attenuation of bottom-up prediction errors (Pecina et al., 2014). Learning theories have underlined the role of prior expectations in the placebo effect; particularly placebo analgesics, where the qualitative experience of pain is overridden by the prediction of pain relief (Rief and Petrie, 2016; Colloca and Benedetti, 2009; De Pascalis et al., 2002; Morton et al., 2014; Wager et al., 2011; Craggs et al., 2007; Atlas et al., 2010; Schenk et al., 2017). Functional imaging studies have identified the neural correlates of changes in the precision of peripheral prediction errors, personality, endogenous opioid system engagement and anticipatory changes that scaffold the effects of placebo hypoalgesia, particularly prefrontal suppression of prediction error processing in the ventral striatum.

Recently, studies have started to look at the role of prediction error traits (Parvaz et al., 2015; Gu et al., 2015b) and interoception separately as markers for substance abuse and relapse (Marhe and Franken, 2014). Using a within subjects placebo design, Gu and colleagues used a

computational model of mesolimbic dopamine systems. They found that prior beliefs about a cigarette's nicotinic content modulated striatum responses to reward prediction errors, evidencing how beliefs can override a potent neuroactive compound, such as nicotine (Gu et al., 2015b). The above studies have explored prediction error traits and interoception in isolation. Integrating these aspects of interoceptive inference therefore remains an outstanding and potentially important challenge.

Under the active inference, anxiogenic traits, such as catastrophizing or somatic hypervigilance can be viewed in terms of the aberrant precision of top-down predictions or bottom-up prediction errors respectively. Therefore, it is possible that anxious individuals may exhibit greater interoceptive accuracy (as measured by heartbeat tracking paradigms) (Schandry, 1981; Ainley et al., 2016), on the view that these individuals assign too much precision to ascending interoceptive prediction errors; *i.e.*, a failure to attenuate ascending interoceptive prediction errors may be associated with undue attention. In line with this, Cornwell and colleagues recently provided evidence that anxiogenic stimuli rebalance feedforward signalling induced by unpredicted sensory input (Cornwell et al., 2017). Specifically, dynamic causal modelling suggests that anxiety-related, hypervigilant responses are best explained by increased postsynaptic gain and modulation of feedforward pathways within a temporo-frontal network. On the other hand, heartbeat tracking tasks 'require' individuals to pay attention to their heartbeat. There is no reason to assume that this reflects a pathological tendency in their habitual interoceptive inferences and indeed the data on anxiety, depression and other psychopathologies are mixed in this respect. On this view, interoceptive accuracy may be a measure of greater flexibility in ascribing precision to interoceptive signals: see (Ainley et al., 2016) for discussion.

In contrast, clinically depressed subjects have diminished interoceptive accuracy (Pollatos et al., 2009; Dunn et al., 2010; Hyett et al., 2015). Recently, reduced resting state connectivity between attentional and interoceptive networks has been found in melancholia (Hyett et al., 2015), offering an explanation for the impoverished interoception and somatic ideation in these patients. These findings suggest that investigating somatic attention and awareness in anxiety and depression may offer targets for behavioural or pharmaceutical treatment strategies. In particular, an interesting clinical question is whether 'normalising' interoceptive precision can affect affective symptomatology. Furthermore, the focus on synaptic gain in the encoding of precision (and its attenuation) speaks to quantifying pathophysiology in terms of effective connectivity; specifically, the intrinsic excitability of neuronal sources in the interoceptive hierarchy. See (Gu et al., 2015a) for an exemplar study that used dynamic causal modelling to look at the intrinsic excitability of the anterior insular, using an empathy for pain task in normal subjects and autistic spectrum disorder.

In hypochondriasis and somatisation disorders, patients report somatic hypervigilance and interoceptive sensitivity (Barsky, 1992; Rief et al., 1998; Ludewig et al., 2005; Anderson and Hope, 2009), indicating aberrant interoceptive precision. Pareés and colleagues (Pareés et al., 2014) report loss of sensory attenuation and a related diminished sense of agency, which is offered as an explanation for functional movement disorder (FMD). This misattribution of agency – regarding voluntary movement – results in FMD patients experiencing the intent to move and actual movement as being simultaneous. Recently, we have applied this paradigm to functional syncopal (fainting) behaviour; *i.e.*, apparent syncope (loss of postural tone and unresponsiveness) during normal blood pressure and heart rate indices that would not cause cerebral hypoperfusion and subsequent loss of consciousness (Mathias et al., 2000). We identified two subgroups that experienced functional syncope during clinical autonomic assessment (Owens et al., 2018b). The first had no undiagnosed form of dysautonomia but a prevalence (41%) of psychiatric illness, presenting as a typical conversion disorder group. The second had no psychiatric illness but autonomic testing revealed undiagnosed postural tachycardia syndrome

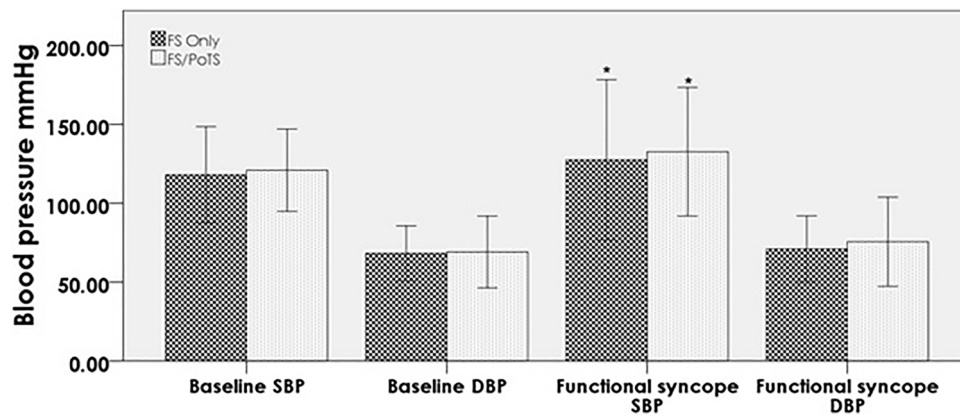


Fig. 4. Baseline and functional syncope episode blood pressure data in the functional syncope only (FS only) and functional syncope/postural tachycardia syndrome (FS/PoTS) cohorts.

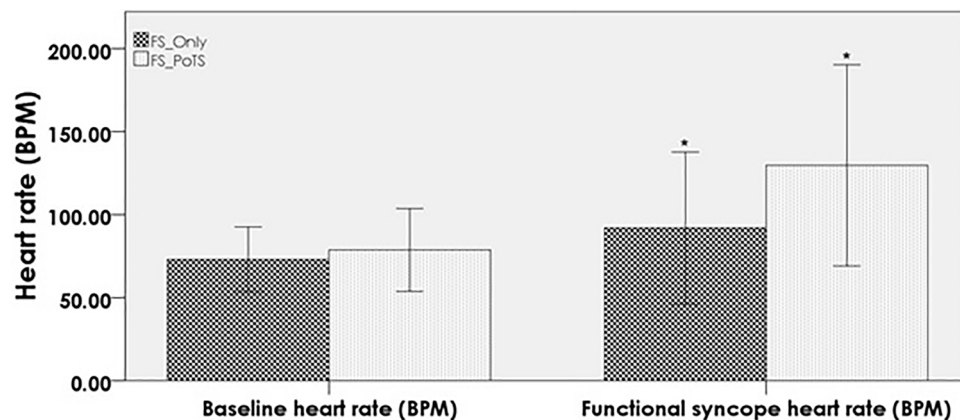


Fig. 5. Baseline and functional syncope episode heart rate data in the functional syncope only (FS only) and functional syncope/postural tachycardia syndrome (FS/PoTS) cohorts.

during orthostatic (upright posture) manoeuvres. Neither group were hypotensive during functional syncopal behaviour episodes (Fig. 4). However, the postural tachycardia syndrome group were typically tachycardic (Fig. 5) during functional syncopal behaviour episodes, which occurred almost entirely during orthostatic stress; i.e., whilst symptomatic with (undiagnosed) orthostatic tachycardia. Some individuals may therefore be prone to impaired sensory attenuation if in a state of undiagnosed sympathoexcitation. One might suppose that the apparent loss of postural tone may reflect a failure to modulate the precision of interoceptive prediction errors during undiagnosed posture-related tachycardia (due to baroreflex dysfunction). This provides a potential explanation for functional syncope in the functional syncope/postural tachycardia syndrome subgroup (Owens et al., 2015b).

Interoceptive inference offers a new and mechanistic perspective on basic and clinical homeostatic issues. For example, with the context-specific knowledge that polite society generally prefers us to micturate in private, ascending lower urinary tract information reaches the brain via the PAG before relaying to the thalamus and hypothalamus, both of which send bladder-related interoceptive signals to the dACC, AIC and IPFC (Griffiths, 2015). If the decision is made not to void, then the medial prefrontal cortex (mPFC) inhibits the PAG. If it is decided that voiding should occur, the mPFC disinhibits the PAG, which activates the pontine micturition centre (PMC). The PMC then engages sacral autonomic efferents to relax the urethral sphincter and contract the detrusor (Griffiths et al., 2005; Fowler et al., 2008). This model provides a nice example of how context-specific information about the environment is inferred (from a Bayesian perspective) to mediate and contextualise autonomic and behavioural homeostatic outputs. Crucially, this rests, under interoceptive inference, on properly

contextualising (i.e., predicting) the precision or gain of interoceptive prediction errors that underwrite homeostatic or allostatic behaviour.

Interoceptive inference proposes that interoceptive predictions and prediction errors can be suppressed by modifying predictions or demarcating these predictions as reference points for autonomically mediated reflexes (Ondobaka et al., 2015a). As with the urinary or cardiovascular systems, thermoregulation can be modelled within the active inference framework. Hypothermia and hyperthermia represent profound deviations from thermostasis; with increasingly complex endocrine, autonomic and behavioural homeostatic reflexes engaged as one's core temperature rises or falls from its homeostatic set point of 37°C (i.e., as prediction error increases) (Collins, 2013). During this process, thermoceptive prediction errors will have greater precision on subsequent central signalling, as glutamatergic, cool-sensitive neurons synapse with GABAergic interneurons in the median preoptic area to initiate thermoregulatory autonomic or motor reflexes (Morrison and Nakamura, 2011). Depending on whether temperature must be increased or decreased, the processing of thermoregulatory prediction error can also result in the inhibition of action potentials in warm-sensitive neurons of the medial preoptic subnucleus, which mediates autonomic control of cutaneous vasoconstriction as well as motor control of shivering and thermogenic brown adipose tissue (BAT) (Benzinger, 1969; Flouris and Cheung, 2009). As interoceptive inference dictates efferent homeostatic changes, BAT neuromodulators, such as glutamate, serotonin and vesicular glutamate transporter 3 will be released to control BAT sympathetic outflow and thermogenesis (Morrison and Nakamura, 2011; Dimicco and Zaretsky, 2007).

Under interoceptive inference, descending predictions can only elicit autonomic responses if the ascending prediction error has been

Table 1

Overview of how interoceptive inference may subjugate autonomic reflexes, as measured by high frequency (HF-HRV) and low frequency heart rate variability (LF-HRV). Correlations between cardiac interoceptive measures and autonomic cardiac control were found in healthy controls whilst supine and orthostatic intolerance patient groups during increased interoceptive prediction error (head-up tilt). Interoceptive accuracy is an objective interoceptive measure gained from the subject's performance during a heartbeat tracking task. Interoceptive sensibility represents subjective confidence in one's own interoceptive accuracy. Interoceptive awareness is a metacognitive measure of the degree to which objective interoceptive accuracy relates to interoceptive sensibility.

Interoceptive inference correlations	Supine HRV	Head-up tilt HRV
Healthy controls	Interoceptive sensibility/LF-HRV ($r = 0.816$, $p = .001$) Interoceptive sensibility/HF-HRV ($r = 0.676$, $p = .002$)	
Postural tachycardia syndrome		Interoceptive awareness/HF-HRV ($r = -0.457$, $p = .043$)
Vasovagal syncope		Interoceptive awareness/HF-HRV ($r = -0.658$, $p = .015$)

attenuated. Without this functional change in gain, prediction errors would lead to revised predictions rather than action (Adams et al., 2013). We recently examined the relationship between measures of cardiac interoception and autonomic cardiac control in healthy controls and patients with forms of cardiovascular dysautonomia defined by baroreflex dysfunction (the postural tachycardia syndrome and vasovagal syncope) to (i) seek empirical support for interoceptive inference and (ii) delineate if this relationship was sensitive to increased interoceptive prediction error in patients during head-up tilt/symptom provocation (Owens et al., 2016, 2018c). Compared to controls, interoceptive accuracy (as measured using a heartbeat tracking task) was reduced in both postural tachycardia syndrome and vasovagal syncope groups. Healthy controls' interoceptive sensibility (subjective confidence in interoceptive accuracy) positively correlated with low-and-high frequency heart rate variability (HRV) whilst supine (Table 1). Conversely, both the postural tachycardia syndrome and vasovagal syncope groups' interoceptive awareness (a metacognitive measure of the degree to which objective interoceptive accuracy relates to interoceptive sensibility) negatively correlated with high-frequency HRV during head-up tilt. Our pilot study offers initial empirical evidence for interoceptive inference and supports our previous findings (Owens et al., 2017b) that postural tachycardia syndrome and vasovagal syncope cohorts share a central pathophysiology underlying interoceptive deficits expressed across distinct cardiovascular autonomic pathophysiology. From a predictive coding perspective, postural tachycardia syndrome and vasovagal syncope patients' data indicates a failure to attenuate/modulate ascending interoceptive prediction errors, reinforced by the concomitant failure to engage autonomic reflexes during head-up tilt. Our findings also define how both central and autonomic processes are ultimately implicated in dysautonomia.

Activation of the right AIC is positively correlated with interoceptive accuracy in healthy controls during heartbeat perception paradigms, with the right insula making inferences about internal bodily states, that can be accessed during conscious interoception (Critchley et al., 2004). The AIC has 2 major roles in interoceptive inference: (i) integrating top-down predictions from high-level cortical regions with bottom-up prediction error and (ii) cascading descending predictions that are a reference point for autonomic mediation of homeostasis (Ondobaka et al., 2015a; Pezzulo et al., 2015). This functional architecture accounts for the recent findings that the degree of damage to the anterior insula is positively correlated with acquired alexithymia levels (Hogeveen et al., 2016), reflecting the interoceptive contribution to inference about emotional states (Damasio, 1999). The AIC contains a significant number of 'von Economo neurons', which are large bipolar, spindle-shaped projection neurons (Seeley et al., 2012). Von Economo neurons are prevalent in humans and are mainly situated in layer Vb of the ACC and the frontoinsula cortex (i.e., the junction of AIC and posterior orbitofrontal cortex) and are specifically associated with interoception (Allman et al., 2011). In comparison to controls, autism spectrum disorder subjects have a significantly greater ratio of von Economo neurons to pyramidal neurons (Santos et al., 2011), which may be of particularly relevance to the common interoceptive sensitivity reported in autism spectrum disorder (Quattrocki and

Friston, 2014; Garfinkel et al., 2016).

The primary motor cortex (M1) is predominantly comprised of agranular neurons and issues motor predictions to the spinal cord to engage motor responses and reflexes (Shipp, 2005). M1 simultaneously sends somatosensory predictions to S1 to model the sensory consequences of the predicted action. The predictions propagated to S1 are efferent copies of motor predictions or commands. We have found that (Adams et al., 2013; Shipp et al., 2013) S1 also attenuates sensory gain during self-initiated movement; thereby reducing prediction error signalling to M1, which receives little direct ascending sensory input. This means predictions descending via M1 to the spinal cord are relatively immune to correction by prediction error. This makes sense if we consider elementary movements are executed in a largely open loop fashion. However, S1 generates predictions about sensory afferent signals that are probabilistic and continuously updated by prediction errors, and changes in the gain of S1 responses are linked to both predictability and attention-driven modulation of felt pain (Fardo et al., 2017).

The functionality of this interoceptive hierarchy can be seen in studies of oesophageal stimulation (Binkofski et al., 1998), where mild stimulation activates secondary somatosensory cortex. Then, as stimulus intensity escalates, interoceptive inference engages primary somatosensory, bilateral insula, ACC and right premotor structures. These results may reflect how escalating interoceptive-to-nociceptive input augments the precision of ascending prediction errors, with subsequent activation of the somatosensory network. If we consider the aberrant interoceptive precision of anxious individuals, this sort of finding may shed light on the fine detail of the neural correlates of irritable bowel syndrome (Blomhoff et al., 2001); particularly in consideration of autonomic (e.g., postprandial) stressors that may augment interoceptive prediction errors in anxious subjects (Van Oudenhove et al., 2016).

3. Viscero-sensory integration, interoceptive self-inference and metacognitive deficits

Reflecting the function of the most central or highest level of the interoceptive hierarchy, metacognitive ability for conscious introspection is frequently disrupted in a variety of psychopathological disorders (Bliksted et al., 2016). Such metacognitive failures; for example, in the case of addiction and posttraumatic stress disorder (PTSD) have been linked to altered arousal (Balconi et al., 2014; Lysaker et al., 2015) that is often highly specific and independent of first-order perceptual or cognitive deficits. Although metacognition has traditionally been cast in terms of signal detection theory as depending solely on the feed-forward recollection of decision-related evidence (Lau and Rosenthal, 2011; Kiani and Shadlen, 2009), recent advances suggest that conscious self-reflection may be better considered as a form of 'interoceptive self-inference', in which hierarchically deep, supramodal predictions of expected precision (or representational stability) enforce interactions between subjective confidence in the interoceptive and exteroceptive domains. For example, we have recently shown that unexpected changes in autonomic arousal reverse the biasing impact of sensory noise (or precision) on subjective confidence, independently of decision

accuracy (Allen et al., 2016b). In a pharmacological follow-up study, we further demonstrated that noradrenaline blockade *via* the beta-adrenoceptor antagonist, propranolol, specifically improves metacognition for perception (Hauser et al., 2018). More generally, confidence for exteroceptive judgements is linked to heart rate increases (Allen et al., 2016b). This is consistent with the hypothesis that metacognition reflects interoceptive-self inference, which not only models the quality of ascending sensory inputs, but also their regulation by the ascending and descending visceromotor processes reviewed above. In this case, metacognitive beliefs are better cast as ‘experiential predictions’ (e.g., I expect to see an apple with high precision and I expect to ‘feel’ good about it), rather than the output of a strictly feedforward sensory process. This view suggests that maladaptive interoception may cause adjustments in metacognitive beliefs and first-order perception, ultimately resulting in disorders such as functional and chronic pain, and social anxiety (in which neutral social stimuli are evaluated as threatening): see also (Stephan et al., 2016). Likewise, deficits in perceptual ability may result in an alteration in autonomic tone, leading to maladaptive decision-making and systematically biased confidence. Collectively, this view motivates empirical investigations into the possibility of a domain-general neural mechanism linking interoceptive and metacognitive inference, raising the importance of measuring visceral-sensory and cognitive deficits using both first and second-order (metacognitive) measures.

4. Conclusions

A decade ago, it was proposed that interoceptive prediction errors could be a bottom-up source of anxiety. Predictive coding models, as assumed under the FEP, propose the brain must recognise the likely cause(s) of afferent sensory input at any given time to support adaptive responses *via* probabilistic (Bayesian) inference. This review provides a framework and supportive evidence suggesting that interoceptive inference can elucidate autonomic control of peripheral effector organs, cognitive-affective function, motor control, consciousness and dissociative symptoms. Insights into the neuroanatomy, neurochemistry, neurophysiology and psychophysiology of active inference, precision and precision-weighting are now beginning to suggest how interoceptive signals inform predictions about the state of the body. This review suggests that interoceptive prediction errors can not only be a bottom-up source of anxiety but may also drive autonomic, metacognitive, motor homeostatic and allostatic systems. A key theme that emerges from this treatment is the role of neuromodulation and synaptic gain control in contextualising the use of ascending prediction errors for interoception and autonomic reflexes respectively – and how subtle deficits in the attenuation of ascending prediction errors can lead to pernicious and diverse pathology.

Acknowledgments

We would like to thank all those who participated in and supported our studies. APO was funded by a research fellowship from the Autonomic Charitable Trust and the St Mary’s Development Trust.

References

Adams, R.A., Shipp, S., Friston, K.J., 2013. Predictions not commands: active inference in the motor system. *Brain Struct. Funct.* 218 (3), 611–643.

Ainley, V., Apps, M.A., Fotopoulou, A., Tsakiris, M., 2016. ‘bodily precision’: a predictive coding account of individual differences in interoceptive accuracy. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 371 (1708).

Allen, M., Friston, K.J., 2016. From cognitivism to autopoiesis: towards a computational framework for the embodied mind. *Synthese* 1–24.

Allen, M., Fardo, F., Dietz, M.J., Hillebrandt, H., Friston, K.J., Rees, G., et al., 2016a. Anterior insula coordinates hierarchical processing of tactile mismatch responses. *NeuroImage* 127, 34–43.

Allen, M., Frank, D., Schwarzkopf, D.S., Fardo, F., Winston, J.S., Hauser, T.U., et al., 2016b. Unexpected arousal modulates the influence of sensory noise on confidence.

eLife 5.

Allman, J.M., Tetreault, N.A., Hakeem, A.Y., Manaye, K.F., Semendeferi, K., Erwin, J.M., et al., 2011. The von Economo neurons in the frontoinsula and anterior cingulate cortex. *Ann. N. Y. Acad. Sci.* 1225, 59–71.

Anderson, E.R., Hope, D.A., 2009. The relationship among social phobia, objective and perceived physiological reactivity, and anxiety sensitivity in an adolescent population. *J. Anxiety Disord.* 23, 18–26.

Atlas, L.Y., Bolger, N., Lindquist, M.A., Wager, T.D., 2010. Brain mediators of predictive cue effects on perceived pain. *J. Neurosci.* 30, 12964–12977.

Auksztulewicz, R., Friston, K., 2015. Attentional enhancement of auditory mismatch responses: a DCM/MEG study. *Cereb. Cortex* 25 (11), 4273–4283.

Aziz, Q., Andersson, J.L., Valind, S., Sundin, A., Hamdy, S., Jones, A.K., et al., 1997. Identification of human brain loci processing esophageal sensation using positron emission tomography. *Gastroenterology* 113 (1), 50–59.

Balconi, M., Finocchiaro, R., Campanella, S., 2014. Reward sensitivity, decisional bias, and metacognitive deficits in cocaine drug addiction. *J. Addict. Med.* 8 (6), 399–406.

Barrett, L.F., Simmons, W.K., 2015. Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* 16 (7), 419–429.

Barsky, A.J., 1992. Amplification, somatization, and the somatoform disorders. *Psychosomatics* 33 (1), 28–34.

Bastos, A.M., Litvak, V., Moran, R., Bosman, C.A., Fries, P., Friston, K.J., 2015. A DCM study of spectral asymmetries in feedforward and feedback connections between visual areas V1 and V4 in the monkey. *NeuroImage* 108, 460–475.

Bauer, M., Stenner, M.P., Friston, K.J., Dolan, R.J., 2014. Attentional modulation of alpha/beta and gamma oscillations reflect functionally distinct processes. *J. Neurosci.* 34 (48), 16117–16125.

Benarroch, E.E., 1993. The central autonomic network: functional organization, dysfunction, and perspective. *Mayo Clin. Proc.* 68 (10), 988–1001.

Benzinger, T.H., 1969. Heat regulation: homeostasis of central temperature in man. *Physiol. Rev.* 49 (4), 671–759.

Binkowski, F., Schnitzler, A., Enck, P., Frieling, T., Posse, S., Seitz, R.J., et al., 1998. Somatic and limbic cortex activation in esophageal distention: a functional magnetic resonance imaging study. *Ann. Neurol.* 44 (5), 811–815.

Blakemore, S.J., Frith, C.D., Wolpert, D.M., 1999. Spatio-temporal prediction modulates the perception of self-produced stimuli. *J. Cogn. Neurosci.* 11 (5), 551–559.

Bliksted, V., Ubukata, S., Koelkebeck, K., 2016. Discriminating autism spectrum disorders from schizophrenia by investigation of mental state attribution on an on-line mentalizing task: a review and meta-analysis. *Schizophr. Res.* 171 (1–3), 16–26.

Blomhoff, S., Spetsten, S., Jacobsen, M.B., Malt, U.F., 2001. Phobic anxiety changes the function of brain-gut axis in irritable bowel syndrome. *Psychosom. Med.* 63 (6), 959–965.

Brown, H., Friston, K., 2012. Dynamic causal modelling of precision and synaptic gain in visual perception - an EEG study. *Neuroimage* 63 (1), 223–231.

Brown, H., Adams, R.A., Parees, I., Edwards, M., Friston, K., 2013. Active inference, sensory attenuation and illusions. *Cogn. Process.* 14 (4), 411–427.

Calejesan, A.A., Kim, S.J., Zhuo, M., 2000. Descending facilitatory modulation of a behavioral nociceptive response by stimulation in the adult rat anterior cingulate cortex. *Eur. J. Pain* 4 (1), 83–96.

Collins, K.J., 2013. Temperature regulation and the autonomic nervous system. In: Mathias, C.J., BR (Eds.), *Autonomic Failure: a Textbook of Clinical Disorders of the Autonomic Nervous System*, 5th ed. Oxford University Press, Oxford p. 247–58.

Colloca, L., Benedetti, F., 2009. Placebo analgesia induced by social observational learning. *Pain* 144 (1–2), 28–34.

Cornwell, B.R., Garrido, M.L., Overstreet, C., Pine, D.S., Grillon, C., 2017. The unpredictable brain under threat: a neurocomputational account of anxious hypervigilance. *Biol. Psychiatry* 82 (6), 447–454.

Craggs, J.G., Price, D.D., Verne, G.N., Perlstein, W.M., Robinson, M.M., 2007. Functional brain interactions that serve cognitive-affective processing during pain and placebo analgesia. *NeuroImage* 38 (4), 720–772.

Craig, A.D., 2002. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3 (8), 655–666.

Critchley, H.D., Corfield, D.R., Chandler, M.P., Mathias, C.J., Dolan, R.J., 2000a. Cerebral correlates of autonomic cardiovascular arousal: a functional neuroimaging investigation in humans. *J. Physiol.* 523 (Pt 1), 259–270.

Critchley, H.D., Elliott, R., Mathias, C.J., Dolan, R.J., 2000b. Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. *J. Neurosci.* 20 (8), 3033–3040.

Critchley, H.D., Mathias, C.J., Josephs, O., O’Doherty, J., Zanini, S., Dewar, B.K., et al., 2003. Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain: J. Neurol.* 126 (Pt 10), 2139–2152.

Critchley, H.D., Wiens, S., Rotshtein, P., Ohman, A., Dolan, R.J., 2004. Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7 (2), 189–195.

Dalton, K.M., Kalin, N.H., Grist, T.M., Davidson, R.J., 2005. Neural-cardiac coupling in threat-evoked anxiety. *J. Cogn. Neurosci.* 17 (6), 969–980.

Damasio, A.R., 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Harcourt Brace, New York.

Damasio, A., Carvalho, G.B., 2013. The nature of feelings: evolutionary and neurobiological origins. *Nat. Rev. Neurosci.* 14 (2), 143–152.

Damasio, A.R., Grabowski, T.J., Bechara, A., Damasio, H., Ponto, L.L., Parvizi, J., et al., 2000. Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat. Neurosci.* 3 (10), 1049–1056.

Dampney, R.A., Coleman, M.J., Fontes, M.A., Hirooka, Y., Horiuchi, J., Li, Y.W., et al., 2002. Central mechanisms underlying short- and long-term regulation of the cardiovascular system. *Clin. Exp. Pharmacol. Physiol.* 29 (4), 261–268.

Dampney, R.A., Furlong, T.M., Horiuchi, J., Iigaya, K., 2013. Role of dorsolateral periaqueductal grey in the coordinated regulation of cardiovascular and respiratory

- function. *Autonom. Neurosci.: Basic Clin.* 175 (1–2). <http://dx.doi.org/10.1016/j.autneu.2012.12.008>. Epub 2013 Jan 19.
- De Pascalis, V., Chiardadia, C., Carotenuto, E., 2002. The contribution of suggestibility and expectation to placebo analgesia phenomenon in an experimental setting. *Pain* 96 (3), 393–402.
- Dimicco, J.A., Zaretsky, D.V., 2007. The dorsomedial hypothalamus: a new player in thermoregulation. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 292 (1), R47–R63.
- Dunn, B.D., Stefanovitch, I., Evans, D., Oliver, C., Hawkins, A., Dalgleish, T., 2010. Can you feel the beat? Interoceptive awareness is an interactive function of anxiety- and depression-specific symptom dimensions. *Behav. Res. Ther.* 48 (11), 1133–1138.
- Fardo, F., Allen, M., Jegindo, E.M., Angrilli, A., Roepstorff, A., 2015. Neurocognitive evidence for mental imagery-driven hypoalgesic and hyperalgesic pain regulation. *NeuroImage* 120, 350–361.
- Fardo, F., Aukstulewicz, R., Allen, M., Dietz, M.J., Roepstorff, A., Friston, K.J., 2017. Expectation violation and attention to pain jointly modulate neural gain in somatosensory cortex. *NeuroImage* 153, 109–121.
- Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1 (1), 1–47.
- FitzGerald, T.H., Moran, R.J., Friston, K.J., Dolan, R.J., 2015. Precision and neuronal dynamics in the human posterior parietal cortex during evidence accumulation. *NeuroImage* 107, 219–228.
- Flouris, A.D., Cheung, S.S., 2009. Human conscious response to thermal input is adjusted to changes in mean body temperature. *Br. J. Sports Med.* 43 (3), 199–203.
- Flynn, F.G., 1999. Anatomy of the insula functional and clinical correlates. *Aphasiology* 13 (1), 55–78.
- Fogelson, N., Litvak, V., Peled, A., Fernandez-del-Olmo, M., Friston, K., 2014. The functional anatomy of schizophrenia: a dynamic causal modeling study of predictive coding. *Schizophr. Res.* 158 (1–3), 204–212.
- Fowler, C.J., Griffiths, D., de Groat, W.C., 2008. The neural control of micturition. *Nat. Rev. Neurosci.* 9 (6), 453–466.
- Friston, K., 2008. Hierarchical models in the brain. *PLoS Comput. Biol.* 4 (11), e1000211.
- Friston, K., 2009. The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13 (7), 293–301.
- Friston, K., 2010. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11 (2), 127–138.
- Garfinkel, S.N., Tiley, C., O’Keefe, S., Harrison, N.A., Seth, A.K., Critchley, H.D., 2016. Discrepancies between dimensions of interoception in autism: implications for emotion and anxiety. *Biol. Psychol.* 114, 117–126.
- Gray, M.A., Beacher, F.D., Minati, L., Nagai, Y., Kemp, A.H., Harrison, N.A., et al., 2012. Emotional appraisal is influenced by cardiac afferent information. *Emotion* 12, 180–191.
- Griffiths, D., 2015. Neural control of micturition in humans: a working model. *Nat. Rev. Urol.* 12 (12), 695–705.
- Griffiths, D., Derbyshire, S., Stenger, A., Resnick, N., 2005. Brain control of normal and overactive bladder. *J. Urol.* 174 (5), 1862–1867.
- Gu, X., Eilam-Stock, T., Zhou, T., Anagnostou, E., Kolevzon, A., Soorya, L., et al., 2015a. Autonomic and brain responses associated with empathy deficits in autism spectrum disorder. *Hum. Brain Mapp.* 36 (9), 3323–3338.
- Gu, X., Lohrenz, T., Salas, R., Baldwin, P.R., Soltani, A., Kirk, U., et al., 2015b. Belief about nicotine selectively modulates value and reward prediction error signals in smokers. *Proc. Natl. Acad. Sci. U. S. A.* 112 (8), 2539–2544.
- Harper, R.M., Bandler, R., Spriggs, D., Alger, J.R., 2000. Lateralized and widespread brain activation during transient blood pressure elevation revealed by magnetic resonance imaging. *J. Comp. Neurol.* 417 (2), 195–204.
- Hauser, T.U., Allen, M., Purg, N., Moutoussis, M., Rees, G., Dolan, R.J., 2018. Noradrenaline blockade specifically enhances metacognitive performance. *eLife* In press.
- Hogeveen, J., Bird, G., Chau, A., Krueger, F., Grafman, J., 2016. Acquired alexithymia following damage to the anterior insula. *Neuropsychologia* 82, 142–148.
- Hohwy, J., 2016. The self-evidencing brain. *Noûs* 50 (2), 259–285.
- Hughes, G., Denantis, A., Wszak, F., 2013. Mechanisms of intentional binding and sensory attenuation: the role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychol. Bull.* 139 (1), 133–151.
- Hyett, M.P., Breakspear, M.J., Friston, K.J., Guo, C.C., Parker, G.B., 2015. Disrupted effective connectivity of cortical systems supporting attention and interoception in melancholia. *JAMA Psychiatry* 72 (4), 350–358.
- Janig, W., Habler, H.J., 2003. Neurophysiological analysis of target-related sympathetic pathways—from animal to human: similarities and differences. *Acta Physiol. Scand.* 177 (3), 255–274.
- Kanai, R., Komura, Y., Shipp, S., Friston, K., 2015. Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 370 (1668).
- Kiani, R., Shadlen, M.N., 2009. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324 (5928), 759–764.
- Kimmerly, D.S., O’Leary, D.D., Menon, R.S., Gati, J.S., Shoemaker, J.K., 2005. Cortical regions associated with autonomic cardiovascular regulation during lower body negative pressure in humans. *J. Physiol.* 569 (Pt 1), 331–345.
- King, A.B., Menon, R.S., Hachinski, V., Cechetto, D.F., 1999. Human forebrain activation by visceral stimuli. *J. Comp. Neurol.* 413 (4), 572–582.
- Klein, T.A., Ullsperger, M., Danielmeier, C., 2013. Error awareness and the insula: links to neurological and psychiatric diseases. *Front. Hum. Neurosci.* 7, 14.
- Lange, C.G., James, W., 1922. *The Emotions* Baltimore. Williams & Wilkins Company.
- Lau, H., Rosenthal, D., 2011. Empirical support for higher-order theories of conscious awareness. *Trends Cogn. Sci.* 15 (8), 365–373.
- Levine, J.D., Gordon, N.C., Fields, H.L., 1978. The mechanism of placebo analgesia. *Lancet* 2 (8091), 654–657.
- Ludewig, S., Geyer, M.A., Ramseier, M., Vollenweider, F.X., Rechsteiner, E., Cattapan-Ludewig, K., 2005. Information-processing deficits and cognitive dysfunction in panic disorder. *J. Psychiatry Neurosci.: Jpn.* 30 (1), 37–43.
- Lysaker, P.H., Dimaggio, G., Wickett-Curtis, A., Kukla, M., Luedtke, B., Vohs, J., et al., 2015. Deficits in metacognitive capacity are related to subjective distress and heightened levels of hyperarousal symptoms in adults with posttraumatic stress disorder. *J. Trauma Dissoc.* 16 (4), 384–398.
- Marhe, R., Franken, I., 2014. Error-related brain activity as a biomarker for cocaine relapse. *Neuropsychopharmacology* 39 (1), 241.
- Mathias, C.J., Deguchi, K., Bleasdale-Barr, K., Smith, S., 2000. Familial vasovagal syncope and pseudosyncope: observations in a case with both natural and adopted siblings. *Clin. Auton. Res.* 10 (1), 43–45.
- Medford, N., Critchley, H.D., 2010. Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. *Brain Struct. Funct.* 214 (5–6), 535–549.
- Moran, R.J., Campo, P., Symmonds, M., Stephan, K.E., Dolan, R.J., Friston, K.J., 2013. Free energy, precision and learning: the role of cholinergic neuromodulation. *J. Neurosci.* 33 (19), 8227–8236.
- Morrison, S.F., Nakamura, K., 2011. Central neural pathways for thermoregulation. *Front. Biosci.* 16, 74–104.
- Morton, D.L., El-Deredy, W., Jones, A.K., 2014. Placebo analgesia: cognition or perception. *Handb. Exp. Pharmacol.* 225, 71–80.
- Mundel, T., Bunn, S.J., Hooper, P.L., Jones, D.A., 2007. The effects of face cooling during hyperthermic exercise in man: evidence for an integrated thermal, neuroendocrine and behavioural response. *Exp. Physiol.* 92 (1), 187–195.
- Nakamura, M., Yoda, T., Crawshaw, L.L., Yasuhara, S., Saito, Y., Kasuga, M., et al., 2008. Regional differences in temperature sensation and thermal comfort in humans. *J. Appl. Physiol.* 105 (6), 1897–1906.
- Ondobaka, S., Kilner, J., Friston, K., 2015a. The Role of Interoceptive Inference in Theory of Mind. *Brain and Cognition*.
- Ondobaka, S., de Lange, F.P., Wittmann, M., Frith, C.D., Bekkering, H., 2015b. Interplay between conceptual expectations and movement predictions underlies action understanding. *Cereb. Cortex* 25 (9), 2566–2573.
- Oppenheimer, S.M., Gelb, A., Girvin, J.P., Hachinski, V.C., 1992. Cardiovascular effects of human insular cortex stimulation. *Neurology* 42 (9), 1727–1732.
- Oppenheimer, S.M., Kedem, G., Martin, W.M., 1996. Left-insular cortex lesions perturb cardiac autonomic tone in humans. *Clin. Auton. Res.* 6 (3), 131–140.
- Owens, A.P., Low, D.A., Critchley, H.D., Mathias, C.J., 2015a. Supine & orthostatic orienting responses to postural interoception in postural tachycardia syndrome (PoTS) & autonomically (Neurally) syncope (AMS): insights into brain-body integration. *Auton. Neurosci.: Basic Clin.* 39–40 Stresa, Italy: Elsevier.
- Owens, A.P., Low, D.A., Critchley, H.D., Mathias, C.J., 2015b. Intermittent autonomic disorders and emotion: a two-way Street? *Auton. Neurosci.: Basic Clin.* 192, 136.
- Owens, A.P., Friston, K.J., Low, D.A., Mathias, C.J., Critchley, H.D., 2016. Investigating the relationship between cardiac interoception and heart rate variability via interoceptive (active) inference. *Clin. Auton. Res.* 26, 315.
- Owens, A.P., Low, D.A., Iodice, V., Mathias, C.J., Critchley, H.D., 2017a. Emotion and the autonomic nervous system - a two-way street: insights from affective, autonomic and dissociative disorders. In: Stein, John (Ed.), *Reference Module in Neuroscience and Biobehavioral Psychology*. Elsevier/9780128093245, pp. 1–15.
- Owens, A.P., Low, D.A., Iodice, V., Critchley, H.D., Mathias, C.J., 2017b. The genesis and presentation of anxiety in disorders of autonomic overexcitation. *Auton. Neurosci.: Basic Clin.* 203, 81–87.
- Owens, A.P., Friston, K.J., Low, D.A., Mathias, C.J., Critchley, H.D., 2018a. Emotional orienting during interoceptive threat in orthostatic intolerance: Dysautonomic contributions to psychological symptomatology in the postural tachycardia syndrome and vasovagal syncope. *Auton. Neurosci.: Basic Clin.* <http://dx.doi.org/10.1016/j.autneu.2018.01.004>. pii: S1566-0702(17)30255-4. [Epub ahead of print].
- Owens, A.P., Low, D.A., Iodice, V., Critchley, H.D., Mathias, C.J., 2018b. Functional syncope behaviour as a maladaptive response to undiagnosed orthostatic intolerance: elucidating medically unexplained symptoms in the postural tachycardia syndrome. *Clin. Auton. Res. Under Rev.*, In preparation.
- Owens, A.P., Friston, K.J., Low, D.A., Mathias, C.J., Critchley, H.D., 2018c. Investigating the relationship between cardiac interoception and autonomic cardiac control using a predictive coding framework. *Auton. Neurosci.: Basic Clin.* <http://dx.doi.org/10.1016/j.autneu.2018.01.001>. pii: S1566-0702(17)30255-2.
- Parees, I., Brown, H., Nuruki, A., Adams, R.A., Davare, M., Bhatia, K.P., et al., 2014. Loss of sensory attenuation in patients with functional (psychogenic) movement disorders. *Brain* 137 (Pt 11), 2916–2921.
- Park, H.D., Correia, S., Ducorps, A., Tallon-Baudry, C., 2014. Spontaneous fluctuations in neural responses to heartbeats predict visual detection. *Nat. Neurosci.* 17 (4), 612–618.
- Parvaz, M.A., Konova, A.B., Proudfit, G.H., Dunning, J.P., Malaker, P., Moeller, S.J., et al., 2015. Impaired neural response to negative prediction errors in cocaine addiction. *J. Neurosci.* 35 (5), 1872–1879.
- Paulus, M.P., Stein, M.B., 2006. An insular view of anxiety. *Biol. Psychiatry* 60 (4), 383–387.
- Pavlov, I., 1927. *Conditioned Reflexes*. Oxford University Press, Oxford.
- Pavlov, I.P., 1960. *Conditioned Reflexes*. Dover Publications, New York.
- Pecina, M., Stohler, C.S., Zubieta, J.K., 2014. Neurobiology of placebo effects: expectations or learning? *Soc. Cogn. Affect. Neurosci.* 9 (7), 1013–1021.
- Pezzulo, G., Rigoli, F., Friston, K., 2015. Active inference, homeostatic regulation and adaptive behavioural control. *Prog. Neurobiol.* 134, 17–35.
- Pollatos, O., Traut-Mattausch, E., Schandry, R., 2009. Differential effects of anxiety and depression on interoceptive accuracy. *Depression Anxiety* 26 (2), 167–173.
- Quattrocki, E., Friston, K., 2014. Autism, oxytocin and interoception. *Neurosci. Biobehav. Rev.* 47c, 410–430.

- Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2 (1), 79–87.
- Rief, W., Petrie, K.J., 2016. Can psychological expectation models be adapted for placebo research? *Front. Psychol.* 28 (7), 1876.
- Rief, W., Hiller, W., Margraf, J., 1998. Cognitive aspects of hypochondriasis and the somatization syndrome. *J. Abnorm. Psychol.* 107 (4), 587–595.
- Santos, M., Uppal, N., Butti, C., Wicinski, B., Schmeidler, J., Giannakopoulos, P., et al., 2011. Von Economo neurons in autism: a stereologic study of the frontoinsula cortex in children. *Brain Res.* 1380, 206–217.
- Saper, C.B., 2002. The central autonomic nervous system: conscious visceral perception and autonomic pattern generation. *Ann. Rev. Neurosci.* 25, 433–469.
- Schandry, R., 1981. Heart beat perception and emotional experience. *Psychophysiology* 18 (4), 483–488.
- Schenk, L.A., Sprenger, C., Onat, S., Colloca, L., Buchel, C., 2017. Suppression of striatal prediction errors by the prefrontal cortex in placebo hypoalgesia. *J. Neurosci.* 37 (40), 9715–9723.
- Schlader, Z.J., Prange, H.D., Mickleborough, T.D., Stager, J.M., 2009. Characteristics of the control of human thermoregulatory behavior. *Physiol. Behav.* 98 (5), 557–562.
- Seeley, W.W., Merkle, F.T., Gaus, S.E., Craig, A.D., Allman, J.M., Hof, P.R., 2012. Distinctive neurons of the anterior cingulate and frontoinsula cortex: a historical perspective. *Cereb. Cortex* 22 (2), 245–250.
- Seth, A.K., Critchley, H.D., 2013. Extending predictive processing to the body: emotion as interoceptive inference. *Behav. Brain Sci.* 36 (3), 227–228.
- Seth, A.K., Suzuki, K., Critchley, H.D., 2011. An interoceptive predictive coding model of conscious presence. *Front. Psychol.* 2, 395.
- Shipp, S., 2005. The importance of being agranular: a comparative account of visual and motor cortex. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 360 (1456), 797–814.
- Shipp, S., 2016. Neural elements for predictive coding. *Front. Psychol.* 7, 1792.
- Shipp, S., Adams, R.A., Friston, K.J., 2013. Reflections on agranular architecture: predictive coding in the motor cortex. *Trends Neurosci.* 36 (12), 706–716.
- Singer, T., Critchley, H.D., Preusschoff, K., 2009. A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* 13 (8), 334–340.
- Skinner, J.E., 1988. Regulation of cardiac vulnerability by the frontal cortex: a new concept of Cannon's cerebral defense mechanism. In: Gailbraith, G.C., Kietzman, M.L., Donchin, E. (Eds.), *Neurophysiology and Psychophysiology: Experimental and Clinical Applications*. Lawrence Erlbaum, Hillsdale, NJ p. 68–80.
- Stephan, K.E., Manjaly, Z.M., Mathys, C.D., Weber, L.A., Paliwal, S., Gard, T., et al., 2016. Allostatic self-efficacy: a metacognitive theory of dyshomeostasis-induced fatigue and depression. *Front. Hum. Neurosci.* 10, 550.
- Tattersall, A.J., Hahn, A.G., Martin, D.T., Febbraio, M.A., 2000. Effects of heat stress on physiological responses and exercise performance in elite cyclists. *J. Sci. Med. Sport/Sports Med. Aust.* 3 (2), 186–193.
- Tucker, R., Marle, T., Lambert, E.V., Noakes, T.D., 2006. The rate of heat storage mediates an anticipatory reduction in exercise intensity during cycling at a fixed rating of perceived exertion. *J. Physiol.* 574 (Pt 3), 905–915.
- Van Oudenhove, L., Tornblom, H., Storsrud, S., Tack, J., Simren, M., 2016. Depression and somatization are associated with increased postprandial symptoms in patients with irritable bowel syndrome. *Gastroenterology* 150 (4), 866–874.
- Vossel, S., Mathys, C., Stephan, K.E., Friston, K.J., 2015. Cortical coupling reflects bayesian belief updating in the deployment of spatial attention. *J. Neurosci.* 35 (33), 11532–11542.
- Wager, T.D., Atlas, L.Y., Leotti, L.A., Rilling, J.K., 2011. Predicting individual differences in placebo analgesia: contributions of brain activity during pain anticipation and pain experience. *J. Neurosci.* 31 (2), 439–452.
- Wiech, K., Lin, C.S., Brodersen, K.H., Bingel, U., Ploner, M., Tracey, I., 2010. Anterior insula integrates information about salience into perceptual decisions about pain. *J. Neurosci.* 30 (48), 16324–16331.
- Wiese, W., 2017. Action is enabled by systematic misrepresentations. *Erkenntnis* 82 (6), 1233–1252.
- Wurtz, R.H., 2008. Neuronal mechanisms of visual stability. *Vis. Res.* 48 (20), 2070–2089.