The neuroscience of interoception in health and disease

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Ann NY Acad Sci</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>annals-1734-011</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Review</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>26-Feb-2018</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Quadt, Lisa; Brighton &amp; Sussex Medical School, Neuroscience Critchley, Hugo; Brighton &amp; Sussex Medical School, Neuroscience; Sackler Centre for Consciousness Science, Psychiatry Garfinkel, Sarah; Brighton &amp; Sussex Medical School, Neuroscience; Sackler Centre for Consciousness Science, Psychiatry</td>
</tr>
<tr>
<td>Keywords:</td>
<td>interoception, health, mental health, predictive processing</td>
</tr>
</tbody>
</table>
The neurobiology of interoception in health and disease

Lisa Quadt¹, Hugo D. Critchley¹,², and Sarah N. Garfinkel¹,²

¹Psychiatry, Department of Neuroscience, Brighton and Sussex Medical School (BSMS), University of Sussex, Brighton, UK
²Sackler Centre for Consciousness Science, University of Sussex, Brighton, UK

Corresponding author:
Lisa Quadt
Trafford Centre
University of Sussex
Falmer
BN1 9RY
L.Quadt@bsms.ac.uk

Keywords
Interoception; health; mental health; predictive processing; autism; anxiety; depression; eating disorders

Abstract
Interoception is the sensing of internal bodily sensations. Interoception is an umbrella term that encompasses: 1) the afferent (body-to-brain) signalling through distinct neural and humoral channels; 2) the neural encoding, representation, and integration of this information concerning internal bodily state; 3) the influence of such information on other perceptions, cognitions and behaviours; and, 4) the psychological expression of these representations as consciously accessible physical sensations and feelings. Interoceptive mechanisms ensure physiological health through the cerebral coordination of homeostatic reflexes and allostatic responses that include motivational behaviours and associated affective and emotional feelings. Furthermore, the conscious, unitary, sense of self in time and space may be grounded on the primacy and lifelong continuity of interoception.

Body-to-brain-interactions influence physical and mental wellbeing. Consequently, systematic investigation of how individual differences, and within-individual changes, in interoceptive processing can contribute to mechanistic understanding of physical and psychological disorders. We present a neurobiological overview of interoception and describe how interoceptive impairments at different levels relate to specific physical and mental health conditions, including sickness and fatigue, depression, eating disorders, autism, and anxiety. We frame these findings in an interoceptive predictive processing framework and highlight potential new avenues for treatments.
1. Introduction

A fundamentally selfish responsibility of the brain is to keep itself, with the rest of the body, alive. The brain coordinates the regulation of vital inner processes, including blood pressure, digestion and breathing, by flexibly reacting to external and internal changes. Interoception refers to the sensing of the internal state of the body, providing the afferent channel of the interplay between body and brain that allows homeostasis (i.e. maintenance of physiological stability) through covert reflexes (e.g. baroreflex), motivational drivers (e.g. hunger and thirst) and explicit bodily sensations (e.g. breathlessness, bladder distension or gastric pain).

Interoception is differentiated by this inwards bodily focus from exteroceptive senses (e.g. vision, audition) that process information about the outer world, and more proximate senses (e.g. proprioception, touch, taste) that use the body to describe the external environment and its relation to it. Interoceptive information is communicated through a set of distinct neural and humoral pathways with different modes of signalling, which the brain represents, integrates and prioritises. How these central representations of the inner body are built and interact is an important focus of interoception research, not least because of the implications for a range of processes and disorders. A comprehensive understanding of cognition, emotion, and overall wellbeing must incorporate an understanding of interoception. The same questions are consequently integral to health neuroscience.

Interoceptive processing has a key role in health and disease, and research is systematically delineating the ways in which brain-body relations can alter a person’s wellbeing.

Interoception involves a relatively restricted set of classes and channels of information (e.g., cardiovascular, gastric, respiratory). These differ with respect to the generation of the signal (organ stretching, mechanoreceptive, chemoreception) and their afferent pathway (neural, humoral). Complexity within interoceptive signalling arises more from the need to parse and integrate of information originating from multiple organs and across wide temporal domains than from the need to differentiate, uniquely characterise and encode complex novel stimuli (even in the generalization of immunological responses). Nevertheless, continuous, dynamic and diverse information about internal bodily function is integrated within shared neural substrates supporting distributed interoceptive representations and associated experiences (feeling states). Together these shape the generative (autonomic or hormonal) control of bodily states and steer adaptive behaviours (e.g. a drop in blood sugar levels leads to foraging). Interoception is not a unitary construct, but can be considered within a conceptual framework encompassing distinct mechanisms and psychological dimensions, characterized for example by processing level (e.g. chemical/neural, behavioural, subjective experiential,
metacognitive), and can be assessed using distinct methodologies (e.g. psychophysiology and brain imaging, behavioural task performance, questionnaire ratings and confidence-accuracy correspondence).\textsuperscript{5, 6} A differentiated, structured view of interoception allows for a fine-grained analysis of concurrent internal processes, how they are represented and communicated to the brain, and how they contribute to health and disease.

One theoretical framework to frame the dynamics and dimensions of interoception is ‘predictive processing’ (PP).\textsuperscript{7-9} The underlying notion is that the brain makes sense of the potentially overwhelming wealth of incoming sensory data by making a ‘best guess’ model of the source of sensory information and tests this model against the afferent data. PP centres on the interplay between bottom-up and top-down processes: Within a neural hierarchy, there is constant communication and interaction between lower- and higher-level processes. Higher-order representations of prior information form the basis of predictions (beliefs or priors) about the expected afferent signal. Such predictions can ‘cancel out’ expected incoming information at lower levels of the hierarchy, permitting ‘prediction errors’ to ascend. These inform and adjust higher-order representations. Cortical neural hierarchies underpin PP models of exteroceptive sensations, e.g. vision. Interoceptive predictive processing (IPP)\textsuperscript{2, 10, 11} describes the hierarchical processing schemes that underlie the interaction between body and brain. For IPP, where informational parameters are arguably more restricted, yet under more direct neural control, cerebral cortex might dominate only at higher-order representational levels.

In this article, we review the dimensional nature of interoception, approaches to their quantification, discuss the neurobiological basis of interoception, and how these findings can be framed within IPP. Importantly, we offer our perspective on the implications for both physical and mental health, and scrutinize the contributing role of interoception to different health conditions. Finally, we suggest how interoception research can further enhance to Health Neuroscience.

2. Dimensions of interoception

Interoception is defined by both its origin within, and reference to, the inner state of the body. This single term generalises communication through multiple distinct physical axes, and representations that unfold at different anatomical and psychological levels, on different timescales. Interoception is a concept that implicitly suggests the integration of different types of sensory information. However, inconsistency within the physiological and psychological literature regarding the definition of interoception, and use of terms such as
interoceptive awareness, led to proposed dimensional frameworks for understanding and studying this set of senses.\textsuperscript{5, 12} Within such a framework, interoception can be described from the physical responses in body and brain representation up to (and beyond) interoceptive metacognitive insight and conscious awareness.

The first dimension refers to the afferent, interoceptive signal that is communicated to the brain from one or more internal organ, which can be measured, for example, by evoked changes in central neural activity, for example as a change in neuroimaging signal or heartbeat evoked potential (HEP).\textsuperscript{13} HEPs refer to a change in neural activity (measured using magnetoencephalography (MEG), electroencephalography (EEG) or intracranial neural recordings) that occurs after a heartbeat. Interestingly, HEP amplitude typically correlates with the ability of an individual to detect and report their heartbeats.\textsuperscript{14}

The second dimension reflects the impact of visceral afferent signals on other forms of central sensory or cognitive processing and behaviours. This level does not necessitate (or preclude) perceptual awareness (i.e. consciousness) of the interoceptive signal or the other processes. Illustrations of this interoceptive dimension are found, for example, in cardiac timing experiments where afferent heartbeat signals impact decisions, emotional processing and memory.\textsuperscript{15-17}

Three ‘psychological’ dimensions refer more directly to the perception of interoceptive signals: Interoceptive accuracy, sensibility and awareness.\textsuperscript{12} These dimensions developed from the use of tests of interoceptive sensitivity/ability, such as heartbeat detection tasks. These tasks are designed to rate individuals according to differences in their ability to sense internal bodily signals, which might account for variation in emotional temperament or psychosomatic vulnerability.\textsuperscript{18} Typically, an interoceptive task requires a participant, at rest, to report ‘felt’ interoceptive sensations (e.g. the timing of a heartbeat): Interoceptive accuracy refers to objective performance on such behavioural tests, e.g. how accurately they perform a heartbeat tracking task.\textsuperscript{19} Next, interoceptive sensibility describes subjective belief about one’s own ability to consciously perceive bodily signals, ascertained via self-report measures such as questionnaires (e.g. body perception questionnaire (BPQ))\textsuperscript{20}, or reflected in their rated confidence in their performance accuracy on an interoceptive task. Since some people think they are good, but in fact are objectively very poor, at reporting bodily sensations (and conversely), this level of conscious insight can be quantified: Metacognitive interoceptive awareness expresses this insight into interoceptive performance aptitude, and can be derived from confidence-accuracy correspondence.\textsuperscript{21} This metacognitive dimension of interoception is a most appropriate use of the word ‘awareness’ in the context of interoception.
A further 'executive' dimension on this interoceptive dimensional framework attempts to capture the degree to which an individual is able to flexibly attend to, and utilize, interoceptive information or can adaptively switch between interoceptive and exteroceptive representations.\(^5\)

3. The neurobiology of interoception

Convergent evidence identifies insular cortex as the brain substrate underpinning higher-order interoceptive representations: for example, left posterior insula cortex is reliably engaged when attention is directed to one’s heartbeat, relative to an exteroceptive focus.\(^22\) Also, anterior insular cortex activity predicts objective performance accuracy on interoceptive tasks. In particular, right anterior insula cortex (AIC) functional reactivity predicts interoceptive accuracy on a heartbeat discrimination task and its volume predicts interoceptive sensibility.\(^1\) The insular cortex is buried between the adjacent frontal and temporal lobes. The architecture of insula changes (including progressive loss of the granule cell layer) from posterior to anterior insular cortex, with other sub-regional differences in cellular organization. Insular cortices are bi-directionally connected to cingulate, prefrontal, parietal, and medial temporal cortices and subcortically to basal ganglia:\(^23\) AIC is strongly connected with anterior cingulate cortex (ACC), arguably forming a functional unit with amygdala and ventromedial/orbitofrontal cortex (VMPFC/OFC), to which they are mutually linked. Posterior insula has stronger reciprocal connections to second somatosensory cortex (SII), and receives direct afferent input from interoceptive thalamus (posterior ventromedial nucleus, which has a lighter corollary projection to anterior cingulate cortex), relaying interoceptive and nociceptive information. Interoceptive information is projected within insula form posterior insula (i.e. primary viscerosensory cortex implicated in primary, objective representations of bodily signals), and rostrally to AIC, which serves to re-represent and integrate interoceptive signals with exteroceptive and motivational information.\(^24\)

The higher-order representation of interoceptive information within AIC and its projection regions underpin consciously accessible feelings that inform emotions and motivate behaviours. This representation also shapes the operational functioning of the brain, as the brain continuously receives and responds to such homeostatic afferent signals. An important aspect of this higher-order representation is the integration across distinct categories of signals that possess distinct temporal response characteristics and encode hormonal, metabolic, thermal, immunological, nociceptive and visceromotor information. This information reaches the brain through humoral and neural pathways.\(^25\) Microglial transduction pathways additionally inform about, and even engage the brain in, inflammatory status, where inflammatory mediators lead to waves in microglial activation that is

http://www.nyas.org/forthcoming
propagated across the brain. However, loss of anatomical specificity, temporal structure and perceptual distinctiveness may be obligatory characteristics of a dynamic higher-order integrative interoceptive representation, from which may emerge an amorphous affective feeling state that is the predictive platform for motivational behaviour, emotional experience, and internal homeostatic control.

Nevertheless, well before insular cortex, conscious access, and affective feeling states, afferent viscerosensory information is processed within subcortical and brainstem regions supporting homeostasis. The nucleus of the solitary tract (NTS) is the main region where visceral neural (spinal laminar 1 and vagus nerve) inputs converge within brainstem, and is of critical importance for the control of physiological state (e.g., blood pressure control). NTS consists of a series of purely sensory nuclei and is organized viscerotopically, where neurons that receive input from distinct organs and types of visceral receptor are in close proximity. This specific organization hints to early integration of viscerosensory signals across related modalities. NTS projects to hypothalamus, ventrolateral medulla and parabrachial nucleus, and through these regions provides a first level of control of hormonal, immune, and autonomic outputs. Chemicals circulating in the blood stream access the brain via specialist circumventricular organs (area postrema, organum vasculosum of laminae terminae and subfornical organ). The humoral information is projected to hypothalamus and NTS, contributing the negative feedback control and cross-modal homeostatic responses mediated through pituitary hormones and the autonomic nervous system.

The NTS receives from spinal visceral afferent neurones with cell bodies in the dorsal root ganglion contain motivational information from cranial nerves, notably the vagus nerve: Viscerosensory inputs with cell bodies in vagus nerve ganglia terminate in the NTS and project onto the pontine parabrachial nucleus, and periaqueductal grey (PAG) before an obligatory relay within posterior ventromedial thalamus. These pre-thalamic midbrain pathways project further to hypothalamus and amygdala, and complement the main viscerosensory thalamocortical projection to insular cortex (and ACC). Nevertheless, all levels of the neuroaxis representing interoceptive information are implicated in the autonomic control of internal physiological state and processes that shape emotions, feelings, behaviour and cognition. Ultimately, the interplay of body and brain depends on bi-directional signal messaging, where higher-level brain regions might influence bodily processes in a top-down manner, and afferent signals influence brain processes from the bottom-up. This complex and dynamic interaction is theoretically captured by an increasingly prominent framework, predictive processing (PP), or, more specifically, interoceptive predictive processing (IPP).
4. Interoceptive Predictive Processing (IPP)

General predictive processing

Predictive processing (PP)\(^7\)\(^,\)\(^8\) is an algorithmic theory about how the brain makes sense of the world and the body it is embedded in. The rationale is that the brain has no direct access to the states of the external world and body, but is instead confronted with an excess of sensory signals. Moreover, each signal has multiple possible causes, so the brain needs to infer the most probable hidden cause of the sensory information it receives.

The external world we live in is full of causal regularities of different spatial and temporal timescales, such as ‘what goes up must go down’.\(^9\) However, these occur alongside noise and irregularities, including unpredictable, surprising events or disturbances in signals. For the brain to use sensory information and to steer the health and behaviour of the organism in an adaptive manner, it must filter out regularities and deal with the noise. PP offers an account about how a neural system finds these regularities. PP suggests that the brain generates a prediction about which input is most likely to arrive next. If this is wrong, a prediction error/mismatch occurs. This error signal can be used in two ways: it can improve and update the model, perhaps generating a perception, (i.e. perceptual inference), or it can lead to a change in behaviour so that the next incoming input fits the prediction better (resulting in action, i.e., active inference). In this way, sensory signals from the outside shape and fundamentally alter predictive representations in the brain; the causal regularities of brain-external matter are ‘folded into’ predictions. This is an important point of PP, as it allows for both the external world and bodily actions to influence the workings of the brain. In other words, it allows for the influence of both environmental, social and cultural factors from the top-down, alongside individual factors, for example genetic dispositions, hormone levels and prior experiences from the bottom-up.

Although PP integrates these brain-external components into its theoretical horizon, it is mainly an account of how the brain works. The basic assumption is that there is a neural and functional hierarchy in the brain that implements generative models. These models are 'generative' because they generate predictions about the most likely state of the level below. The cortical hierarchy implements predictions that range from highly abstract regularities (e.g. ‘what goes up must come down’) at higher levels to basic, concrete sensory properties of incoming signals at lower levels. Higher levels putatively operate at slower timescales, while timescales get faster as one goes down the hierarchy.\(^33\) Along this hierarchical organization, predictions are passed down and compared to the actual state of the level below, all the way down to the sensory input. The discrepancy between prediction and signal is propagated back up the hierarchy, where it is used to change generative models to improve their predictive power. This process of prediction error minimization (PEM) lies at
the heart of PP and is thought to be the brain’s primary task – improving it’s guesses about what is going on outside the skull, so it can steer behaviour in the most efficient way.

**Interoceptive Inference**

Interoceptive inference, or interoceptive predictive processing (IPP) takes up the general PP framework and applies it to describe internal body-brain interactions. Here, high-level predictions about the internal state of the body are generated within cortex (AIC is most strongly implicated) within a neural hierarchy, proximately involving posterior insula. Descending predictions are compared against incoming afferents, creating an error signal that serves to improve predictions and reduce subsequent prediction error through both perceptual inference (change in feeling state) and active inference (autonomic and behavioural response). These generative predictions cascade to earlier levels of control (including brainstem autonomic centres, which operate along similar negative control feedback principles), ultimately serving to keep bodily states within their expected range for adaptive behaviour, thereby keeping the physiological integrity.

The Embodied Predictive Interoceptive Coding (EPIC) model relates IPP and prediction error minimization more specifically to cortical architecture. By analogy to predictive coding within the motor system, EPIC proposes that interoceptive predictions originate in the deep layers of agranular (i.e., less laminar differentiation) visceromotor regions within prefrontal (caudal VMPFC/OFC), anterior / mid cingulate cortices and AIC. Back-projecting predictions are proposed to terminate within the superficial layers of dysgranular and granular cortical columns, where they alter ongoing pattern of activity by changing the firing range of neurons in anticipation of viscerosensory sensory input. These interoceptive inputs ascend from the NTS, parabrachial nucleus, via thalamus to primary dysgranular and granular regions of mid- and posterior insular cortex. There, it is proposed that cortical prediction errors are computed (i.e. difference between predicted and actual signal). The resulting prediction error signal is then projected onto the deep layers of agranular visceromotor cortices, where the prediction originated. At this point the error signal can trigger the generation of new descending predictions that are ultimately expressed as autonomic/visceromotor outputs. This process is interoceptive active inference minimising future prediction error through generating interoceptive inputs that confirm predictions. Alternatively, the error may trigger a reduction of further signal sampling to reduce further prediction error (impacting feeling state). Lastly, another option is that the error signal adjusts the precision of prediction units within visceromotor cortices thereby modulating sensory sampling and viscerosensory input through adjusting the gain on thalamocortical communication.
The EPIC model of interoceptive predictive processing also suggests, in line with the general principle of predictive processing, that interoceptive sensations are largely driven by predictions. This means that the perception of bodily signals is weighted toward mostly top-down, rather than a bottom-up, cortical processes. The perception of bodily sensations is thus determined by predictions that are informed by prior experience and kept in check by actual bodily states. The extent to which these predictions lead to perception also depends on precision-weighting (instantiated at one level as attention) across the interoceptive hierarchy, where precision units reflect both the reliability of prediction and prediction errors to increase or decrease the gain on error signals in order to change predictions. A well-functioning precision-weighting system is paramount for healthy functioning, as will become more obvious in later parts of this paper.

Interoceptive predictions interact with other sensory modalities, projecting onto visual, auditory and somatosensory networks, to provide an embodied representational context for perception cognition and action. This way, interoceptive representations modulate responses across the brain, which serves as a reference for exteroceptive process and enable a dynamic multisensory representation of the body in its environment. What we perceive and how we behave is thus ultimately influenced by interoceptive predictions and is steered towards keeping ourselves alive and well. Agranular cortices, the putative origin of interoceptive predictions, are less constrained by incoming signals from the body, in turn may permit predictions to be abstract and directed towards the future, enabling allostasis in place of the reactive maintenance of homeostasis. IPP therefore encapsulates the flexible interplay between top-down and bottom-up processes that support stable, yet dynamic, internal environment.

In a healthy brain, predictions are informed by prior experience, situational context and state of the system, the comparison between prediction and actual incoming bodily signal, and precision estimation that results in a well-balanced interaction of brain and body. The goal of this complex process is to keep bodily states within a functional that permits flexibly adaptation to both internal changes and external challenges. The interoceptive system balances anticipated demands and deviations, efficiently regulating needs and resources. This process was conceptualized as ‘allostasis’ or ‘predictive regulation’ and is underpins the well-being of body and mind.

5. Interoception and physical health

The processing of interoceptive signals in the brain informs central control processes involved in maintaining physiological integrity. Interoception is tightly related to the predictive control of bodily signals that contribute to a system being able to maintain homeostatic set-
points, and the flexible allostatic regulation of more complex demands. When the system fails to respond to demands in an adaptive manner, or when predictive fluctuations fail to foresee necessary demands, the organism may reach allostatic overload and succumb to sickness and disease.

**Sickness behaviours**

The human immune system communicates immunological and inflammatory states to the brain via interoceptive pathways. The peripheral states of infection and inflammation are transmitted to the brain via vagus nerve pathways, cytokines that circulate humorally, and via immune cells. Responses to these insults include the activation of cardiovascular and gastrointestinal reflexes, the regulation of peripheral immune reactions, and also a stereotyped pattern of responses called ‘sickness behaviours’. These entail fatigue, reduced calorie and fluids intake, social isolation, anhedonia, and fever. Sickness behaviours are thought to facilitate counteracting responses to infection and inflammation by inducing behavioural patterns that reduce bodily strain (e.g. fatigue motivates rest), and risk of additional infection (e.g. social isolation). This narrow repertoire of behaviours is evoked as a response to a wide range of infectious and inflammatory conditions, which suggests that they may form a coordinated general physiological and motivational reaction to a particular type of interoceptive challenge for the protection of the body’s integrity.

Experimentally these mechanisms can be explored by administration of substances that cause a brief spike in inflammation, e.g. typhoid vaccine, infusion of endotoxin, or inhalation of antigens. A neurally-mediated interoceptive pathway, recruiting basal and posterior ventromedial thalamus, and dorsal mid- and posterior insula, is activated after typhoid vaccination. Specific components of sickness behaviour are associated with functional changes within interoceptive brain regions, including mid-insula (fatigue), subgenual cingulate (mood change), and the midbrain substantia nigra (psychomotor slowing). The insula is further implicated in the expression of inflammation-induced subjective experiences of fatigue, malaise and social disconnect. Increase in right anterior insula metabolism tracks the loss of interest in social interaction, while heightened connectivity between anterior insula and mid-cingulate cortex predict subjective malaise and discomfort after induction of inflammation. These findings indicate a role for the insula in mediating the experiential side of sickness behaviours, a hypothesis that is in line with the theoretical proposal and emerging evidence implicating insular cortex in subjective experience of conscious motivational and emotional states arising from interoceptive predictive processing.
The same brain regions that support emotions and affective regulation are thus involved in sickness behaviours (and their origin in IPP), highlighting a connection between inflammation, sickness behaviour and mood disorders.\textsuperscript{44} Changes in motivation are a hallmark of both sickness behaviours and major depressive disorder.\textsuperscript{48} Low motivation to move can be adaptive in the context of physical illness, as it enables energy conservation while prioritizing resources for fighting off inflammation and infection. In the case of prolonged or very severe inflammation, however, these motivational changes can mark the onset of a depressive episode.\textsuperscript{44} Motivational changes ultimately impact processing of reward-stimuli,\textsuperscript{18, 19} correspondingly response to reward outcomes is altered following inflammation. This is reflected on both the neural and behavioural level; reactivity within the ventral striatum, a centre of (predictive) reward processing\textsuperscript{49} is decreased, and both subjective and objective measures of anhedonia (the absence of reactivity to positive stimuli) are increased.\textsuperscript{40} Social withdrawal is another symptom that sickness behaviours and depression share. Not participating in social interaction often leads to feelings of isolation and loneliness, and contributes to the maintenance of depressed mood.\textsuperscript{50} Inflammation, through interoception, thus facilitates processes that underlie and enhance feelings of social isolation; induce feelings of social disconnect,\textsuperscript{51} and impair the processing of social cues.\textsuperscript{52} Taken together, sickness behaviours illustrate how perturbation of internal bodily states impact neural representations, emotional states, and executive behaviours. These reactive patterned responses are mediated via interoceptive pathways that typically support adaptive social emotional and motivational behaviours.

Fatigue

Fatigue is a disorder that is characterized in the ICD-10 as a long-term condition that includes severe and constant feelings of tiredness, trouble concentrating and carrying out daily activities, generalized aches and pains, fever, and sleep disturbances.\textsuperscript{53} It can be part of sickness behaviours, and as such have adaptive effects in that it prioritizes rest to save resources and may facilitate the role of fever in fighting off infections.\textsuperscript{54} Fatigue can also appear on its own as a chronic condition (chronic fatigue syndrome)\textsuperscript{55}, which affects approximately 20% of the general population.\textsuperscript{56} Its prevalence increases to 50%, however, as a symptom in conditions that are associated with a compromised immune system,\textsuperscript{57} such as cancer,\textsuperscript{58} autoimmune diseases like multiple sclerosis,\textsuperscript{59} and fibromyalgia.\textsuperscript{60} Fatigue is strongly associated with depression,\textsuperscript{61} and listed in both DSM-5 and ICD-10 as a core criterion for major depression.\textsuperscript{53, 62}

Fatigue is a multi-dimensional construct that involves both impairment of motor and cognitive processes, and the subjective experience of fatigue.\textsuperscript{63} Research on fatigue emphasises
approaches that associate the condition with peripheral inflammation and its influence on brain structures involved in steering immunological responses. Brain structures involved in fatigue include insula and the frontostriatal network, most notably the ventral striatum. In this context, signals of peripheral inflammation reach the frontostriatal network via immune-to-brain communication pathways that involve activation of microglia. This network underlies response to reward, which supports anticipation and motivation, both of which are reduced in fatigue. An altered frontostriatal network due to inflammation is thus one strong candidate for the neurobiology of fatigue. AIC has been associated with the experiential quality of emotions and feelings, and is thought to play a key role in the experience of fatigue. After the experimental induction of inflammation via typhoid vaccine, fatigue was predicted by altered reactivity within mid- and posterior insula and ACC. This suggests that interoceptive signalling of inflammatory states, and their impact on brain regions that are associated with processing interoceptive input, is an important factor in subjective experience of fatigue and vitality/agency. Newly emerging views on fatigue are turning towards approaches that not only consider the bottom-up effects leading to fatigue, but that also take into account possible top-down influences. Further research is needed to determine if distinct levels of interoceptive processing accuracy are compromised in individuals with high levels of fatigue.

6. Interoception and Mental Health

Interoception research is increasingly demonstrating that, in addition to physical health, the signalling and detection of internal bodily signals is important for mental wellbeing. Interoceptive and emotional processes share underlying neural substrates, and prominent theories of emotion even suggest that emotional feeling states arise through the sensing of bodily signals. Emotional impairments accompany the majority of mental disorders, acting as one potential route linking interoception to mental health.

Depression

Major depressive disorder is associated with affective symptoms such as low mood, and negative cognitions such pervasive negative thoughts and intense feelings of hopelessness. In addition, somatic symptoms including aches and pains, disordered sleep, loss of appetite and fatigue are just as frequent, and occur universally across cultures. Recognition that somatic alterations are an important factor for changes in emotion and cognition has grown over the past decade. Depression is associated with autonomic dysfunction, manifesting as decreased baroreflex sensitivity, reduced phasic skin conductance responses, and reduced heart rate variability. In addition to autonomic alterations, signs of heightened inflammation have been documented in depression.
subset of individuals with depression, cumulative meta-analyses demonstrate raised inflammatory markers, particularly IL-6 and C-reactive protein. Disturbances in brain function are linked to increases in peripheral inflammatory markers, where, for example, reduced functional connectivity of corticostratial reward circuitry is observed in depressed individuals with elevated C-reactive protein.

Impaired interoceptive accuracy may lead to reduced emotional experience, and indeed, ‘feeling nothing’ is often reported by depressed individuals. Healthy controls demonstrate a correlation between accuracy and intensity of experienced emotions, where better accuracy leads to more intense feelings, raising the possibility of a potential impairment in interoceptive accuracy in depression. However, experiments detailing altered patterns of altered interoceptive accuracy associated with depression, present a more complex relationship. The ability to accurately perceive one’s heartbeat is negatively correlated with depression symptoms in healthy controls, an effect only found to manifest when coupled with high anxiety. In an experiment which contrasted interoceptive accuracy across three groups (healthy controls, community sample with moderate depression and a more severely depressed clinical sample), only the moderately depressed sample had significantly impaired interoception. Interestingly, and counter to predictions, the more depressed group displayed levels of interoceptive accuracy that were comparable to the control group, though this effect may have been influenced, in part, by medication status. Increasingly, nuanced investigation of interoceptive behavioural impairments linked to specific clusters of symptoms (e.g. differentiating negative effect from emotional numbness) may reveal clearer associations in depression.

Decreased heartbeat perception accuracy is accompanied by significantly reduced heart beat evoked potential (HEP) amplitudes in depressed individuals. The neurocircuitry underlying attention to visceral interoceptive sensations was assessed in unmedicated individuals with major depressive disorder (MDD) relative to controls. Activity in the dorsal mid-insula as well as a network of brain regions involved in emotion and visceral control, were decreased in the MDD group. Moreover, resting state functional connectivity between the amygdala and the dorsal mid insula cortex was increased in MDD and predictive of depression severity. Together these results suggest that the brain representation of interoceptive focus may be altered in MDD.

From a theoretical approach, IPP (including the EPIC model) provide insight into depressive mechanisms, extending to the hypothesis that structural abnormalities, and dysfunctional metabolism within agranular visceromotor cortices may be underlying causes of depressive
states, particularly when associated with inflammation and sickness behaviours.\textsuperscript{2} Visceromotor cortical dysfunction causes imbalance between demand and response through over-predicting metabolic energy-demands. This may engender overactivity of the hypothalamus-pituitary-adrenal (HPA) axis and thereby increasing levels of pro-inflammatory cytokines,\textsuperscript{89} causing concomitant alterations in the immune and endocrine system.\textsuperscript{90} This aberrant process will compromise dependent coupling of interoceptive predictions and inputs at the thalamocortical level, leading to a speculated increase in interoceptive prediction errors. Down-regulation of these noisy error signals by precision units leaves them less able to influence and inform predictions. To further reduce prediction errors, the interoceptive network is left with two principle options; maintaining the dysfunctional predictions, or generating afferents that match these predictions. The latter is thought to lead to noisier signals, setting them up to still not be able to change predictions. This insensitivity to prediction errors might mean that faulty predictions will maintain metabolic energy demand, until the endocrine and immune system have reached their limit. Depression ensues when the error signals can finally no longer be ignored and must be reduced, enlisting sickness behaviours to conserve energy.\textsuperscript{2} The insensitivity to prediction errors in combination with ever-more demanding predictions is hypothesised to lead to a ‘locked-in’ (attractor state) brain that maintains a vicious cycle of faulty predictions and noisy error signals.\textsuperscript{91} Inefficient energy-regulation may underlie negative affect, biasing the system more towards avoidance behaviours and social withdrawal.\textsuperscript{92} An IPP model of depression (and fatigue) thus connects aberrant allostatic processes to imbalanced affective processing, driving both somatic and experiential emotional symptoms of depression.

**Autism Spectrum Conditions**

Autism spectrum conditions (ASCs) are classified as neurodevelopmental conditions that are associated with stereotypical and restricted behavioural patterns, altered sensory reactivity, and social and emotional impairments.\textsuperscript{93}

Research is currently investigating the nature of interoceptive deficits associated with ASCs. Work in children is divergent, with one study suggesting interoceptive accuracy is intact in children and adolescents (aged 8-17) with ASCs\textsuperscript{94} while a subsequent study found that interoceptive accuracy, ascertained using heartbeat tracking, was markedly impaired in a comparable child and adolescent autistic sample.\textsuperscript{95} Impaired interoceptive accuracy has also been shown in adults with ASCs, demonstrated using the heart beat tracking task, where significantly lower interoceptive accuracy scores were observed relative to a matched control group.\textsuperscript{6} One study, however, demonstrates data to suggest that autism per se does not
necessarily lead to interoceptive impairments, but instead alexithymia, which is highly co-morbid with ASCs, is associated with reduced interoceptive accuracy.\textsuperscript{96} Alexithymia is a subclinical condition characterized by a reduced capacity to detect and identify emotions in oneself and others,\textsuperscript{97} and thus the emotion processing deficits in autism, characterised by high alexithymia, may be the principle driver for interoceptive impairments in ASC. Other studies in non-autistic populations have demonstrated a link between high alexithymia and impairments in interoceptive accuracy.\textsuperscript{98} Together these results suggest that interoceptive accuracy may be impaired in individuals with autism, and that this may be particularly coupled with emotion processing deficits.

In contrast to behavioural performance on interoceptive tests, interoceptive sensibility, assessed via self-report questionnaires, is elevated in adults with ASCs, despite these same individuals demonstrating a relative impairment in interoceptive accuracy.\textsuperscript{6} This is in line with research documenting that interoceptive aptitude ascertained using self-report does not necessarily predict actual performance measures.\textsuperscript{12} Moreover, it suggests that these interoception dimensions may further diverge in clinical populations, with ASC individuals having an overinflated belief in their interoceptive aptitude relative to their performance accuracy. This enlarged discrepancy between objective and subjective interoceptive performance denotes potential poor interoceptive sensory precision in ASCs and is in line with accounts of autism conceptualized as a condition with an imbalance of the precision ascribed to sensory evidence relative to prior beliefs.\textsuperscript{99}

Altered Insula reactivity has been observed in individuals in ASC across a variety of distinct emotion processing tasks, including response inhibition of emotion al stimuli,\textsuperscript{100} processing of bodily expressions,\textsuperscript{101} and the processing of incongruent emotional information.\textsuperscript{102} ASC is also associated with altered intrinsic functional connectivity of anterior and posterior insula regions and specific brain regions involved in emotion and sensory processing.\textsuperscript{103} Together, these results suggest that altered sensory precision marked by reduced interoceptive accuracy underscored by aberrant insula activity and functional connectivity may contribute to emotion processing deficits observed in ASC and alexithymia more generally.

Anxiety Disorders

Anxiety disorders include panic disorder, agoraphobia, social anxiety, generalized anxiety disorder (GAD), and specific phobias.\textsuperscript{62} Investigations into interoceptive alterations in anxiety disorders are mixed, reflecting the diversity of anxiety conditions and also the range of methodological approaches.\textsuperscript{104} Studies have reliably found that interoceptive sensibility, i.e. self-report measures of interoception, are elevated in individuals with a variety of anxiety
related conditions. In accordance with this, interoceptive accuracy is also frequently elevated in individuals with anxiety, indexed by heightened performance on heartbeat perception tests in patients with anxiety and elevated occurrence of trait anxiety symptoms with heightened interoceptive accuracy in non-clinical cohorts. However, a straightforward relationship between elevated interoception in anxiety is challenged by a number of studies which either do not show a relationship, or reveal a reverse relationship, with higher levels of anxiety related to reduced interoceptive accuracy. Recent work partly reconciles these divergent findings, by demonstrating that it is the relationship between subjective and objective measures of interoception which predict anxiety symptomatology (in both an autistic population and in healthy controls). Specifically, individuals with an elevated interoceptive trait prediction error, derived from a propensity to belief one is interoceptively proficient despite relatively poor interoceptive accuracy, had heightened trait anxiety scores. This interoceptive predictive error is potentially consistent with theoretical work that has posited that the pathogenesis of anxiety is related to noisy interoceptive input in combination with noisily amplified self-referential interoceptive predictive belief states.

Eating Disorders

Eating disorders (EDs) are characterized by atypical food intake (e.g., restriction in anorexia nervosa, or binging and purging in bulimia nervosa), and are often accompanied by a distorted body image. Poor interoception has been linked to body image concerns, and a number of empirical findings converge to suggest potential disturbances in the processing of interoceptive signals in individuals with EDs. Interoceptive self-report in this population has been primarily probed using the Eating Disorder Inventory (EDI) which assesses the subjectively reported ability to discriminate sensations of hunger and satiety, and to respond to emotional states. Patients with EDs report impairments in these abilities, which could reflect a generalized deficit in interoceptive processing. Empirical findings support this in part, with studies demonstrating impaired interoceptive accuracy in anorexia nervosa patients relative to matched controls using a heartbeat perception test. Other studies, however, fail to show impaired interoceptive accuracy in anorexia nervosa, and instead document enhanced reported detection of interoceptive sensations.

To date, only few studies have investigated whether interoception is compromised in bulimia nervosa, although it is suggested that interoceptive processing deficits drive the symptoms and associated behaviours in bulimia. One study investigating interoceptive accuracy in woman with a current diagnosis of bulimia nervosa observed no differences in heartbeat tracking task performance when correcting for the presence of co-varying comorbid alexithymia, depressive symptoms and anxiety. In contrast, women who had recovered
from bulimia nervosa (without a prior diagnosis of anorexia nervosa) demonstrated significantly reduced interoceptive accuracy compared to controls.122

Neural representation of bodily state is altered in EDs. During an interoceptive attention task (focusing on the heart, stomach and bladder) individuals with anorexia nervosa display significantly reduced activation in the anterior insula during heart perception, and significantly reduced activation in the dorsal mid-insula during stomach interoception, relative to a matched control group.123 Individuals with anorexia nervosa display reductions in functional connectivity in the thalamo-insula subnetwork, thought to reflect changes in the propagation of sensations that convey homeostatic imbalances.124 Bulimia nervosa is associated with increased gray matter volumes within the ventral anterior insula,125 and binge eating disorder is associated with increased insula activity when viewing food images after an overnight fast.126

Interestingly, altered interoception is not only found in patients who are currently suffering from an eating disorder. Impairments in interoceptive self-report, as measured by the EDI, predicts vulnerability to the development of EDs, as revealed in longitudinal studies.127-129 It is not yet known whether other dimensions of interoception, such as interoceptive accuracy or neural processing of bodily state, would also demonstrate pre-morbid alterations. Nevertheless, interoceptive measures, as least ascertained via self-report, may serve as a marker for ED vulnerability, facilitating potential early intervention.

The exact nature of interoceptive impairment in EDs remains unclear, as it varies across types of eating disorder, and studies often do not take into account comorbidities such as anxiety, depression and alexithymia, which are also associated with aberrant interoception.87 130 Differences in methodology also potentially contribute to further ambiguity, with objective and subjective dimensions of interoception being used interchangeably, and the interoceptive axis (e.g. cardiac vs. gastric) also requiring further differentiation and systematic evaluation. Behavioural, neuroimaging, and psychophysiological studies nonetheless show that several dimensions of interoception are affected in different types of EDs. Further research with terminological and methodological consistency could help to create a more differentiated account of how interoception contributes to, and maybe even predicts, the occurrence of eating disorders.
7. Conclusion

There is increasing evidence that the signalling, sensing and detection of bodily states are implicated in physical and mental wellbeing.\textsuperscript{69, 131} Interoception research contributes an important dimension to Health Neuroscience, by providing powerful explanatory understanding into the dynamic interactions between body, brain and mind that underlie pathophysiological disturbances across physical and mental disorders. Capitalising of strengthening theoretical frameworks, including IPP, further research needs to extend systematic interoceptive investigation across different bodily axes, and include measures of interoception that cover neural signalling, objective behavioural performance, subjective experiences and beliefs, alongside metacognitive measures, to delineate comprehensively interoceptive predictors of specific symptoms. Understanding the precise nature of interoceptive deficits has important clinical implications, as insight into interoceptive mechanisms may reveal new therapeutic targets to promote novel interventions.
References


http://www.nyas.org/forthcoming


70. James, W., *What is an emotion?* Mind, 1884. 9: p. 188-205.


