Interoception beyond homeostasis: affect, cognition and mental health

Manos Tsakiris\textsuperscript{1,†} and Hugo Critchley\textsuperscript{2,3,†}

\textsuperscript{1}Laboratory of Action and Body, Department of Psychology, Royal Holloway, University of London, Egham TW20 0EX, UK
\textsuperscript{2}Psychiatry Group, Department of Neuroscience, Brighton and Sussex Medical School, Universities of Sussex and Brighton, Brighton BN1 9FX, UK
\textsuperscript{3}Sackler Centre for Consciousness Science, University of Sussex, Brighton BN1 9RH, UK

Interoception refers to the sensing of the internal state of one’s body. Interoception is distinct from the processing of sensory information concerning external (non-self) stimuli (e.g. vision, hearing, touch and smell) and is the afferent axis to internal (autonomic and hormonal) physiological control. However, the impact of interoception extends beyond homeostatic/allostatic reflexes: it is proposed to be fundamental to motivation, emotion (affective feelings and behaviours), social cognition and self-awareness. This view is supported by a growing body of experimental evidence that links peripheral physiological states to mental processes. Within this framework, the representation of self is constructed from early development through continuous integrative representation of biological data from the body, to form the basis for those aspects of conscious awareness grounded on the subjective sense of being a unique individual. This theme issue of the Philosophical Transactions of the Royal Society B draws together state-of-the-art knowledge concerning theoretical, experimental and clinical facets of interoception with the emphasis on cognitive and affective neuroscience. The multidisciplinary and cross-disciplinary perspectives represented in this theme issue disseminate and entrench knowledge about interoception across the scientific community and provide a reference for the conceptualization and further study of interoception across behavioural sciences.

1. Introduction

Interoception is the sensory system that communicates the internal state of the body through signals originating from within the visceral organs [1–3]. This encompasses information about the functional state and health of the organs (distension, pressure, motility, tissue damage). A broader and increasingly accepted definition ascribes interoception the role of both sensing and integrating all aspects of the body’s physiological state and motivational needs, from low-level monitoring of blood chemistry, the representations of skin and body temperature and sensations evoked by pleasant interpersonal touch [3]. Interoception, as the sense of the physiological condition of the body, supports homeostatic control and allostatic adaptation, ensuring the stability of the organism [4] and by driving behaviour through feelings such as hunger, thirst and dyspnoea. Moreover, interoceptive signals are increasingly recognized to have a pervasive (as yet incompletely characterized) impact on cognition, influencing attention and perception, guiding decision-making and shaping memory and emotion processing [5,6]. As such, interoception is a complex phenomenon that presents several different dimensions. Interoceptive deficits are now recognized as important factors in the expression and maintenance of many psychiatric and neurological disorders, including anxiety and depression, addiction and anorexia [7]. More radically, interoception has been recently linked to phenomenal consciousness [3,8,9] and body awareness [10,11]. Recent theoretical developments, supported by a timely expansion of the evidence-base, have pushed interoception into the foreground of...
human cognitive neuroscience. Over the last 10 years, there has been a sixfold increase in the number of new publications on the topic of interoception. This increase reflects convergent interdisciplinary appreciation of how interoceptive processes can enhance the accounts of mechanisms that underpin affective and cognitive functions, including the coherent representation of a conscious sense of self. These insights further inform our understanding of emotional psychopathology.

There is now a pressing need to reach consensus regarding the scientific study of interoception. Across physiology, sensory neuroscience, consciousness science, philosophy and medicine, academics are working on interoception without necessarily sharing the same conceptual base or necessarily realizing how their investigations link with the findings and insights of others. This theme issue attempts to address this need by presenting key recent advances that draw together anatomical and physiological knowledge, theoretical perspectives, testable models of neural information processing (e.g. predictive coding) with experimental and clinical neuroscientific findings. Our key aim is to compile and present a compelling body of empirical evidence and neurobiological insights into how information about the internal states of the body are integrated with thoughts, feelings and behaviour. We draw together current advances across different research areas and the related disciplines of psychology, neuroscience, neurology, psychiatry and psychophysiology. These questions are deeply interconnected and inherently interdisciplinary in nature. Building on these questions, the following sections highlight the main contributions made by each paper to this theme issue.

2. Interoceptive processing: awareness and cognition

The first area of research in the theme issue considers the effects of interoceptive processing and individual differences in interoceptive ability on higher cognitive functions, including self-consciousness and memory.

Babo-Rebelo et al. [12] focus on the role that the neural monitoring of internal bodily signals might play in self-consciousness. Drawing on intracranial electroencephalography and magnetoencephalography recordings, the authors provide fresh insight into the way brain systems, notably the ‘default mode network’ and right anterior insula, support the encoding of self-related thoughts through the monitoring of cardiac function. Their data extend previous research showing that neural responses to heartbeats within the default mode network encode two distinct self-dimensions, the ‘I’ and the ‘me’. The authors propose that rather than representational convergence on a given brain area, a unifying mechanism underlying both the cognitive and bodily self is the neural monitoring of visceral organs.

García-Cordero et al. [13] examine whether distinct dimensions of interoception, i.e. accuracy, learning and awareness, rely on specialized neural subregions distributed throughout the brain within an ‘interoceptive network’. They focus on healthy participants and on three patient groups that provide complementary lesion-deficit models through neurodegeneration and focal brain damage (i.e. behavioural variant frontotemporal dementia, Alzheimer’s disease and frontoinsular stroke). The lesion model was combined with measures of high-density EEG, as well as structural (MRI) and functional connectivity imaging (fMRI). On testing, each of the three patient groups manifested deficits in interoceptive accuracy (here, accuracy for detecting heartbeat signals at rest together with deficits in the heart-evoked potential). Importantly, interoceptive learning was specifically impaired in patients with Alzheimer’s disease, implicating mesial temporal lobe memory networks in this skill. Measures of interoceptive awareness, i.e. metacognitive correspondence between subjective and objective measures of interoception, on the other hand, showed that both frontotemporal dementia and Alzheimer’s disease patient groups overestimated their performance with reduced insight. Together these findings highlight how damage to specific hubs and related connections within a frontotemporinsular network differentially compromises dimensions of interoception.

Umeda et al. [14] focus on prospective memory (PM), the cognitive ability to retain offline and re-invoke at an appropriate point, memories for future intentions. They expand on past work that link the performance on tests of PM to physiological arousal associated with stress, suggesting autonomous ‘somatic markers’ might trigger PM retrieval, an effect that is mediated by individual differences in interoceptive accuracy. In their study, participants with higher PM task performance showed a greater heart rate increase when PM targets were presented. Furthermore, participants with higher interoceptive accuracy showed better PM task performance. Taken together, these results highlight the dependence of an important aspect of higher cognition on dynamic and trait expressions of interoception.

Across sensorimotor neuroscience, predictive coding increasingly provides a powerful framework for understanding brain function from the neuronal level to behaviour. Within this theme issue, three papers apply predictive coding models explicitly to the understanding of interoception [15–17]. Ainley et al.‘s [15] application of a predictive coding/free energy framework to interoception explores its theoretical plausibility against published evidence. They focus on how precision across the interoceptive system accounts for individual differences in interoceptive performance: high interoceptive accuracy is pro- posed to reflect the capacity to adjust priors (descending expectations) to minimize ascending prediction errors through attentional deployment (alongside reflexive physiological responses). This model usefully accounts for phenomenological and related observations of experimental studies employing heartbeat detection tasks, and ultimately has implications for understanding clinical disorders in terms of interoceptive dysfunction, through its attention-dependent contribution to self-hood and cognition (including decision-making).

3. Interoception and affect

Long-standing debates about interoception centre on its contribution to affect and emotion processing [18–20]. Within this theme issue, the second research area draws together novel empirical data and theoretical perspectives to examine contributions of interoceptive processing to affect, emotions and pain perception.

Strigo & Craig [21] provide an integrated account of interoception and its neuroanatomical substrates in a paper that draws together a refreshed appraisal of the compelling neuroanatomical evidence for a broad definition of
interoception as the ‘physiological sense of the condition of the body’. This definition, which includes thermal sensation and affiliative sensual touch is motivated by Craig’s characterization of the laminar 1 spinothalamic pathway and its projections within the primate brain. In this paper, this evidence is extended by exciting new data regarding the parallel vagus nerve afferent brain pathways, assumed to converge in brainstem with ascending laminar 1 information, and now identified to hold a characteristic thalamic topography. Critically, Craig’s neuroanatomical model of interoceptive brain hypotheses a basis for lateralized affective process grounded on distinct sympathetic/spinal and parasympathetic/vagus nerve anatomy. Strigo and Craig here provide empirical support for this proposal in a neuroimaging study that combines slow breathing manipulation with positive versus negative affective challenge (processing emotional pictures). They show predicted co-lateralization of affect/valence and bodily representation (cardiorespiratory arousal) at the level of viscerosensory insula.

Krahe et al. [22] explore a boundary of interoception: the perception of affective touch conveyed by unmyelinated C-fibres. Such sensation is encompassed within Craig’s broadened model of interoception, a view supported both by their anatomical organization through laminar 1 spinothalamic tract and by clinical observations [23]. Here, the authors measure electroencephalographic potentials associated with pain and tactile somatosensation to differentiate affective from non-affective attributes of these sensations, particularly the pleasant ‘stroking’ sensation of C-tactile fibre stimulation. Importantly, the authors demonstrate the link between these sensory responses and affect, as crystallized in individual differences in attachment style. These observations reinforce the evidence for including C-tactile sensations within a broader definition of interoception and their corresponding contributions to motivational, affective and affiliative behaviour.

Seth & Friston [16] present interoceptive inference as a powerful conceptual tool that contextualizes interoception in terms of recent influential models of predictive coding. Predictive coding has emerged as a prominent unitary theoretical account of brain function that explains cortical processes underlying perception, action and (more recently) interoception [24,25]. According to the theory, incoming sensory data are compared with internal models, i.e. the brain’s probabilistic ‘prediction’ (best guess) about the causes that affect the organism’s nervous system. If predictions and data are not compatible, then ‘prediction errors’ arise. However, organisms must maintain their bodies within a narrow range of desirable states, and therefore prediction errors must be minimized. Seth and Friston’s notion of interoceptive inference means that bodily states are regulated by autonomic reflexes that are enslaved by descending predictions from deep generative models of our internal and external milieu. This re-conceptualization illuminates several issues in cognitive and clinical neuroscience with implications for experiences of selfhood and emotion. The authors provide a complementary narrative to the papers by Ainley et al. [15], and Barrett et al. [17] contained in this theme issue.

4. Clinical implications of interoceptive processing

Ultimately, research on interoception is fundamentally motivated by important clinical questions, as interoception is increasingly recognized as relevant to health and the understanding and evaluation of brain (psychiatric and neurological) as well as bodily disorders. Clinical advances are represented within this section of the theme issue, though many clinical areas informed by the study of interoception, including medically unexplained symptoms [26], somatization [27] and addiction [28], are equally important though not explicitly covered within this issue.

Garfinkel et al. [29] examine interoceptive dimensions across respiratory and cardiac axes and define how these relate to the expression of anxiety symptoms. They dissociate cardiac and respiratory measures of interoceptive accuracy (i.e. objective task performance), yet demonstrate a positive relationship between cardiac and respiratory measures of interoceptive awareness (i.e. ‘higher’ metacognitive insights into own interoceptive ability which transfers across these senses). Importantly, poor respiratory accuracy was associated with heightened anxiety score, whereas good metacognitive awareness for cardiac interoception was observed to be protective against anxiety. These findings highlight how distinct psychological dimensions of interoception relate to anxiety across respiratory and cardiac axes. This has important implications for the optimization of anxiolytic effects associated with specific therapeutic interventions.

Influential theoretical models predict that the impact of interoception varies as a function of the degree to which an individual is sensitive to internal signals. The dimension model of Garfinkel [30] has arguably provided a step-change in approaches to the study of such interoceptive individual abilities, by differentiating objective interoceptive ‘accuracy’ (quantifiable through discriminatory performance on interoceptive tasks), from subjective accounts of interoceptive experience (‘sensibility’ confidence, belief, questionnaire ratings) and from interoceptive ‘metacognitive awareness’, reflecting insight quantifiable as the degree to which objective and subjective measures correspond.

Here, Mehling [31] provides a powerful psychological critique of this dimensional model, including the terminology (e.g. the task-constrained use of interoceptive metacognitive awareness), and presents a refined approach to the measurement of subjective interoceptive ‘sensibility’ to overcome the typically negative (anxiety-related) bias associated with many interoception questionnaires and to capture positive interoceptive phenomenology, including those associated with meditative experience. The refinements are encapsulated within the multi-dimensional assessment of the interoceptive awareness questionnaire, which is gaining clinical and research utility.

As noted above (see also [15,16]), predictive coding models of brain function applied to interoception are able to powerfully account for associated aspects of perception, behaviour and neural organization. In an elegant paper, Barrett et al. [17] draw together these conceptual, experimentally determined and neuroanatomical features within a predictive coding model that emphasizes active inference (i.e. behaviours and viscero-motor activity motivated to minimize sensory interoceptive prediction error), as the basis of allostatic control linked to affect. Perturbation of this system is proposed to be fundamental to depression. Using anatomical insights concerning the functional organization of cortical architecture, the authors provide a compelling account of how allostatic brain mechanisms are locked into maladaptive control patterns. Evidence for this is implicit in neuroimaging studies of depressive psychopathology. This work points to the utility of interventional strategies that embody this level of interoceptive knowledge.
5. Measuring and modulating interoception

Despite recent advances [29,30,32], there is consensus that interoception research must develop a more grounded measurement model, applicable across disciplines, that permits a fuller characterization of the links between different interoceptive systems, if interoception is to maintain its growing status within psychology. This ambitious aim necessitates wide-ranging, dedicated and systematic theoretical and methodological enquiries into (i) horizontal relations across interoceptive modalities, (ii) hierarchical relations in interoceptive processing, and (iii) causal relations between objective interoception, subjective experience and metacognitive awareness, instead of mere correlations between them.

Classic approaches to quantifying interoceptive sensitivity focus on four systems: cardiovascular, respiratory, gastrointestinal and urogenital [2,33]. A fundamental question concerns the interrelation of awareness across different interoceptive systems. Some studies report significant correlations between cardiac and respiratory sensitivity [34], and cardiac and gastric sensitivity [35], whereas others do not ([36], also [29]). Interestingly, much neuroimaging evidence does suggest commonalities, suggesting convergent representation across interoceptive modalities within specific brain regions, notably insula cortex [3]. However, there is an implicit assumption that cardiac detection (perhaps by virtue of the utility of associated tests) represents an indicator of ‘general’ interoceptive ability constraining the wider generalization of findings. Still, psychological research into interoceptive ability has focused mainly on cardiac detection. However, such tests are acknowledged to have perhaps insurmountable psychometric weaknesses (e.g. explicit measures of cardiac interoceptive accuracy may be confounded by participants’ beliefs about their heart rate and their estimation of the time elapsed during trials [37]). These confounds can generally be mitigated by combining more than one such task, or using appropriate control conditions, as was previously demonstrated [38]. Nevertheless, the question of how different hierarchical levels of interoceptive processing relate to each other remains important for optimizing our methods.

Another major methodological limitation relates to the ability to causally manipulate interoceptive processing. Unlike exteroception, it is typically difficult to maintain close experimental control over inputs into the interoceptive system. In the last section of this theme issue, methodological advances are considered that address hierarchical, horizontal and causal relations in interoception.

Brener & Ring [39] highlight the difficulties in collecting objective evidence of interoception that result from the intended use of non-invasive behavioural measures of accurate perception of heartbeat sensations as a proxy of interoceptive sensitivity. They focus their critique on the two most popular methods for assessing interoceptive sensitivity, heartbeat tracking and the two alternative forced choice methods, and question their psychometric validity, compromised for example by expectations, prior knowledge and practice. To address these limitations, Brener and Ring suggest the use of classical psychophysical approaches as a means of providing unbiased measures of the temporal locations of heartbeat sensations and the precision with which these sensations are detected.

Hassanpour et al. [40] have developed a fine-gained pharmacological methodology to manipulate internal bodily state for the assessment of individual differences in interoceptive ability. Infusing isoproterenol intravenously stimulates cardiorespiratory arousal peripherally in a dose-dependent way, and induces corresponding feeling states (that localize to the chest). They used a functional neuroimaging study in healthy participants to demonstrate the distinct engagement of subregions of the insula during anticipation of cardiorespiratory arousal versus the actual experience of these physiological changes. Intriguingly, this approach dissociates the activation of left and right posterior mid insula, perhaps linked to the mechanisms and control processes described in Craig’s conceptual model of cortical lateralization. Right anterior insula was activated with the perception of physiological arousal, which occurred in all participants at higher doses of isoproterenol. This methodology, with functional neural correlates now mapped in the brain, offers much to the field of human interoception.

Pollatos et al. [41] apply for the first time non-invasive brain stimulation in healthy participants to causally influence dimensions of interoceptive accuracy and awareness. They used repetitive transcranial magnetic stimulation in a continuous theta burst protocol to target the inhibition of specific central locations of the interoceptive network including the right anterior insula and the right somatosensory cortices and assessed the effects on interoceptive processing. Inhibition of anterior insula function resulted in a significant decline in both cardiac and respiratory interoceptive accuracy and in an accompanying decrease in perception confidence, whereas stimulation over somatosensory cortices reduced cardiac interoceptive accuracy and affected perception confidence. These findings pave the way for the development of mechanistic neuropsychological models of causal relations within interoception, and of strategies to test these models.

The eyeblink startle is a protective motor reflex, the magnitude of which can be used psychophysologically to index affective state. Schulz et al. [42] have previously shown that the eyeblink startle is modulated by interoceptive cardiac afferent signals. Here, the authors focus on the respiratory system and for the first time present startle modulation by respiration. Participants were presented with acoustic startle noises during a spontaneous or a paced-breathing condition. The highest startle magnitudes were observed during expiratory phases of respiration, independent of the breathing condition. Afferent signals from slowly adapting phasic pulmonary stretch receptors are implicated in this effect. These findings highlight the potential of startle-reflex methodology as a non-invasive, quantitative index of interoceptive processing, now extended to include respiratory-related afferent signalling.

Schulz [43] presents a meta-analysis of functional neuroimaging (fMRI) studies of interoceptive attentiveness (i.e. focused attention to a particular interoceptive signal for a given time interval) to one’s heartbeat. Nine such studies were submitted to multilevel kernel density analysis to reveal an extended network that encompassed posterior insula, claustrum, medial frontal gyrus, superior temporal gyrus and precentral gyrus. Within insula, a right-hemispheric dominance was observed in association with cardioception, whereas prefrontal neural activity was implicated in top-down attention deployment and processing of feedforward interoceptive information.
6. Conclusion

This theme issue achieves its objectives in drawing together state-of-the-art theoretical accounts with new experimental (and anatomical) evidence and clinical data to provide an enriched knowledge base with respect to interoception and the different methodologies through which it can be shown to influence cognitive and affective processes. However, there are some key messages that emerge from the papers compiled within this theme issue: there still remains issues around broad versus wider definitions of interoception (e.g. can cutaneous touch sensations, even if affectively loaded, really be co-categorized with homeostatic afferents from the deep viscera?). Moreover, what of chemosensing, from olfaction and taste to blood glucose and oxygen sensing? In fact, there may be no hard boundaries to interoception, but equally, internal organs—notably the heart—have clearly been co-categorized with homeostatic afferents from the deep viscera. Beyond homeostasis, and pave the way for new directions in interoceptive research across health and disease.

Competing interests. We declare we have no competing interests.

Funding. M.T. received support from the European Research Council (ERC-2010-StG-262853) under the FP7. H.C. received support from the European Research Council (ERC-2012-AdG-324150) and the Dr Mortimer and Theresa Sackler Foundation via the Sackler Centre for Consciousness Science, University of Sussex.

References

18. James W. 1884 What is an emotion? Mind 9, 188–205.


