Journal of NeuroPhilosophy 2023;2(2):241-275

In"tarot"ception – Primer, Learning Module, Philosophical Foundations and Application

Jesse S. Bettinger

Abstract

Perception operates in general according to two causal processes—*bottom-up* (sensory driven; progressive), and *top-down* (anticipatory; inferential; with signal progression issuing from expectations about future states). Both are essential to consciousness. Like colors, there are, in essence, three primary modes of perception. Exteroception, interoception and proprioception. The interoceptive faculty of perception refers to the moment-to-moment awareness of the body's interior dynamics in concert with regulatory homeostatic and allostatic operations that furnish a core feature of subjective experience, social navigation and sense-of-self. This paper provides an overview of interoception accompanied by a visual diagram that serves as a heuristic learning model for conceptually organizing the materials. This follows with a summary of A.N. Whitehead's 'three modes of perception' that clarifies the connections with interoception to such an extent they may be regarded as philosophical foundations. This is followed by the topics of information integration and biological intuition.

Key Words: Perception, interoception, neurovisceral axis, predictive coding, free-energy principle, Bayesian inference, somatic markers, insula, integrated information

DOI: 10.5281/zenodo.10200160

Introduction

The various modes of awareness afforded by sensory faculties can be conceptually housed under the category of perception. Increasingly it is becoming a philosophical standard to model perception in terms of a synthesis of causal signaling processes characterizable as *bottom-up* (sensorial; responsive) and *top-down* (model-based; anticipatory, inferential). Such dynamics are not unique to perception; rather, the integration of top-down and bottom-up information serves as a general organizational principle for the brain (Dayan *et al.,* 1995; Friston, 2002, 2012; Corbetta *et al.,* 2008) and consciousness (Dehaene *et al.,* 2006; Dehaene & Changeux, 2011; Edlund *et al.,* 2011). As Italian psychiatrist and developer of "integrated information theory," Guilio Tononi, and the German neurophysiologist Christoph Koch explain,

Corresponding author: Jesse S. Bettinger
Address: Center for Process Studies, USA
e-mail ✉ jesse.bettinger@cgu.edu

"the integration of bottom-up input and top-down recurrent information is considered essential for consciousness to occur" (Tononi & Koch, 2008).

Generally speaking, bottom-up processes are causally-progressive, meaning sensory inputs are accumulative, following the arrow-of-time and signal-causality forward from *cause to effect.* Such processes are mostly stimulus-driven and responsive. This approach operates on the implicit claim that “perception follows sensation and therefore bodily feelings originate in the body” (Barrett & Simmons, 2015). Alternatively, top-down processes in perception are mainly regulatory, driven by intrinsic models of probabilistic expectations, imperatives, priorities, values, beliefs, self (identity) and social contexts. In neural dynamics, inferential predictions of possible causes (as well as future states of current sensory signals) issue from models of past experiences—i.e. from *effects to causes.* As Barrett and Simmons explain:

> Rather than neurons simply lying dormant until information arrives via the external sensors of the body (that is, the eyes, ears and taste receptors, among others), the brain anticipates incoming sensory inputs, which it implements as predictions that cascade throughout the cortex. (Barrett & Simmons, 2015)

This additional operation provides an adaptive benefit to speed of response, preparation, survival and social navigation by allowing the organism to anticipate events rather than just respond. Along with these two causal processes, perception can be further categorized into three primary modes (similar to the primary colors) as: *exteroception*, *interoception* and *proprioception* (Ondobaka, Kilner & Friston, 2015).[2] Of these, this paper highlights interoception as the affective sense of the physiological conditions of stimuli originating in the body (Garfinkle & Critchley, 2013). The purpose for providing a simplified, multidimensional overview of interoception is four-fold. a) to reconcile the notion of a two-way, causal operation converging in interoception; b) to provide an accompanying "conceptual-learning model" for interoception; c) to recognize this topic includes precise philosophical foundations; d) to provide a proof-of-principle for a biological basis of "Integrated Information Theory."

Over the coming pages the case will be made that interoception is not only a worthy topic for general knowledge—with practical benefits involving emotional intelligence, intuition, empathy, resilience, health, wellness, motivation—but also an academic learning-locus that brings together a plethora of fields along with new pedagogical strategies and curricular developments for teaching in philosophy, psychology and neuroscience. What is more, interoception appears to possess one of the clearest examples of ready-made

[2] *Exteroception* refers to the classical five senses; *proprioception* refers to external bodily position and movement.

philosophical foundations by way of the modes of perception outlined in 1929 in the writings of A.N. Whitehead. Finally, the topic of interoception may provide a chief exemplar of the type of processing sought by IIT with regard to subjective streams of experience. This paper seeks to parameterize the conceptual landscape of interoception in such a way that this key topic can be easily packaged into a format that facilitates the learning comprehension of empirical topics.

Outline of Topics

The order of topics in this paper proceeds with an introduction and overview of interoception developed in the context of *homeostasis; the internal milieu*, *neurovisceral axis* and *insular dynamics*. This is followed by a description of interoception in the complementary top-down causal operation driven by approximated *Bayesian prediction/precision-coding* and the *free-energy principle*. Accompanying the topics in this paper is a visual model that will prove useful as a learning device. This follows with an overview of A.N. Whitehead's "three modes of perception" that is sufficient to show how a textbook-ready level of philosophical foundations can be precisely and vividly identified when read into the context of interoception. The discussion section contextualizes Tononi's computational model of consciousness as a measure of "integrated information" in the neurobiological setting of "interoceptive information integration in the insula." Following this, a third track of the model is introduced involving biological intuition. The paper closes with a review and final comments.

Interoception (Background)

“[People] have really got a new sense, and found within their world another world, or nest of worlds.”

Ralph Waldo Emerson, 1844

The faculty of *interoception* refers to an affective sense and awareness of the interior physiological conditions and visceral sensations arising from the inner organs and tissues of the body (Craig, 2002), including the detection of subtle changes in bodily states like the muscles, joints and viscera (Dunn & Galton *et al.,* 2010). Interoception/interoceptive awareness entails "a phylogenetically-novel homeostatic sensory afferent pathway in primates, especially [hominids] that provides the basis for a sense of the physiological condition of the body" (Craig, 2002). It is noteworthy that this does not preclude other species from similar expressions. To some degree the underlying architecture is shared across all mammals (Craig, 2003). Rather, it emphasizes the adaptation of species-specific affordances in linguistic, social and emotional processing that follow from more-complex intricacies of social dynamics. As the Damasio's explains, "given that *homo sapiens* developed the largest groups of any primate, they have the most complex social lives and consequently needed the most highly developed and widely ranging set of emotions" (Damasio & Damasio, 2022).

Through the constant cortical rendering of physical feelings, physiological conditions and gut microbial signatures, the interoceptive faculty of perception furnishes an ongoing representation of “the bodily self” (Craig, 2003) linked to multiple effects on higher-level cognitive and psychological phenomena such as *memory formation*; *emotional arousal*; *affective behaviors* (Cameron, 2002); *motivation*; *mood*; *sense of well-being*; *behavioral regulation; emotional intelligence; composure*; *resilience*; *social cognition*; *intuitive decision making* (Damasio, 2003; Mayer, 2011; Verdejo-Garcia *et al.,* 2012); *self-representation* (Schachter & Singer, 1962; Damasio *et al.,* 1991; Lazarus, 1991); *alexithymia* (lack of emotional self-awareness linked to low interoceptive acuity: Ernst *et al.,* 2014); and as a basis for "*embodied feedback theories*" (Damasio, 1994; Dunn & Dalgleish *et al.,* 2010) — reinforcing the role of interoception as an important factor in framing subjective and physiological experience and a sense-of-self.

The relationship between bodily- and emotional- awareness dates back almost 150 years to the critical scholarship of psychologists William James (James, 1884) and Carl Lange (Lange, 1885) who ascribed emotional feelings to internal bodily responses with the proposal that emotions only come to be perceived in awareness if specific bodily processes are simultaneously perceived. In this "embodied" model, emotions are caused by the perception of physiological signals automatically elicited by stimuli (Bear *et al.,* 2020) such that “our feeling of the same changes as they occur *is* the emotion” (James, 1884). On either side of that generation in psychology, two biologists, Claude Bernard (Bernard, 1865) and Walter Cannon (Cannon, 1929) developed important ideas that later proved foundational to the contextualization of the neurovisceral axis; namely, *homeostasis* and the *internal milieu* of the body. For a thorough review of the contributions of Bernard and Cannon, please see (Lee, 2019).

Homeostasis and the Internal Milieu

As far back as the early Greeks, the philosopher Heraclitus (540–480 BC) hypothesized that a static, unchanging state was not the natural human condition, and that the ability to undergo constant change was intrinsic to all things (Clendening, 1942). Around the same time, Empedocles (492-435 BC) postulated that balance and harmony were necessary conditions for the survival of living organisms (Stavros, 2014). These two views align like conceptual harbingers to what in the latter 1800's, Claude Bernard described as the "milieu interieur" and in the early 1900's Walter Cannon "amplified" into the notion of homeostasis, as a basis for the regulation of dynamical bodily systems (Lee, 2019).

Notably, in 1865 Claude Bernard published a landmark paper stating that "the maintenance of the internal environment, the inner environment, surrounding the body's cells, was essential for the life of the organism" (see Goldstein, 2007). Bernard's idea was two-fold; namely, that there exist an identifiable set of physiological parameters defining the normative internal states of the organism; and that the body would seek to maintain the operation of these optimal parameters (Gross, 1998). Inspired by Robin and Bernard's "milieu intérieur," in 1929 Walter Cannon formalized the concept of homeostasis as "a process of synchronized adjustments in the internal environment resulting in the maintenance of specific physiological variables within defined parameters including blood pressure, temperature and pH balance" (Cannon, 1929; see also Buchman, 2002). As Buldeo explains, "both the milieu interieur and homeostasis imply the existence in the body of mechanisms by which the organism can track the moment-to-moment fluctuations of these physiological parameters" (Buldeo, 2015). In this way, homeostasis is an endogenous operation involved in maintaining the internal stability of an organism.

Cannon's key insight was that the body must manage a multiplicity of complex and variable interactions in order to maintain constancy (homeostasis) or to return systems to normal-range functioning (resilience), and that maintaining this balance requires an internal axis of communication in the form of afferent, visceral sensors found in all organs and tissues of the body that "identify deviations from the acceptable ranges, and effectors to return those deviations back within acceptable limits" (Cannon, 1939). That we would have such a regulatory construct makes a lot of sense (for background see Lee, 2019; Billman, 2020). Given the various biological and physiological demands involved in maintaining "the body's immanent autonomic, metabolic and immunological needs" (Barrett & Simmons, 2015), in order to persist in a long-term sustainable fashion requires the organism to restrict itself to an adaptive-subset of 'goldilocks' biophysical states (Friston *et al.,* 2013). In terms of information biology and thermodynamics, the ability of the body to regulate its internal physiology via homeostatic processes exhibits an express proclivity "to resist the second law of thermodynamics […] and thereby prevent states from disbursing over all possible states they could occupy" (Friston *et al.,* 2015). These ideas lay the foundation for interoception as an affective faculty of experiencing the activity associated with an anticipatory and responsive internal milieu.

The Neurovisceral Axis

By 1950, interoception was beginning to be recognized as a faculty of perception distinct from exteroception (Sherrington, 1948). In the last few decades, advances in functional neuroimaging and physiology have made it abundantly clear that the brain and gut are in constant communication, sending the activity of various neurocrine, endocrine, microbial, autonomic and immune signals (e.g. cytokines and *entero-endocrine cells*: Mayer, 2011) encoding sensory information in the gut between cerebral, autonomic and enteric centers (Craig, 2003; Rhee, Pothoulakis & Mayer, 2009; Grenham *et al.,* 2011; Holzer, 2017).

The *neurovisceral* (brain–gut) axis serves as "a ubiquitous information channel used to represent one's body from within" (Craig, 2002). This axis has been meticulously mapped by the German gastroenterologist Emeran Mayer (Mayer, 2011). This information is primarily communicated by small-diameter (C and Aδ) sensory fibers in all organs/tissues reflecting various changes in sensitization of intrinsic states throughout the body like temperature; pain; muscle contractions in blood vessels; hormonal activity; and cytokine inflammation (Chernigovskiy, 1967; Vaitl, 1996; Tajadura-Jimenez & Costantini, 2011). As Lee explains, electron microscopic findings of the ubiquity of nerve fibers on blood vessels, endocrine cells, and elsewhere suggest that "the brain has close access to essentially every somatic cell" (Sterling, 2004; Lee, 2019).

Efferent sensory fibers are classified as *sympathetic* or *parasympathetic* while *afferent* sensory fibers (traveling→brain) are distinguished by *vagal* and *spinal* pathways (Janig, Levine, & Michaelis, 1996). Parasympathetic efferents are predominantly excitatory and mediated by cholinergic neurotransmission of: motor fibers secreting acetylcholine to striated muscles; parasympathetic fibers to involuntary muscles; and to sympathetic ganglia from the central nervous system (Hale, Shekhar, & Lowry, 2012). Alternatively, inhibitory sympathetic efferents utilize similar adrenergic pathways resembling cholinergic pathways, except where norepinephrine is the neurotransmitter instead of acetylcholine (Alosachie *et al.,* 1998; Hale *et al.,* 2012). These elements facilitate compensatory activity essential to survival by "enabling the body to holistically control or maintain this homeostatic milieu interieur" (Buldeo, 2015).

In addition, there are two afferent pathways: (i) "lamina I spinothalamocortical system" (sympathetic output) and *nucleus tractus solitarius* (parasympathetic output) projecting to the posterior insula and cingulate cortex; and (ii) the somatosensory pathway projecting to the primary somatosensory cortex (see Khalsa, 2008; Khalsa *et al.,* 2009). The constant relay of neurovisceral afferent fibers following the *spino-thalamo-cortical* path targets three key areas: the *insular cortex*, *anterior cingulate cortex* and *somatosensory cortex*. Of these primary processing regions, afferent fibers amalgamate into temporal "representational images"[3] in the insula—a highly-integrative brain region linked to activity involving self-awareness, emotional experience, body-ownership (Craig, 2009), plus membership in a 'rich-club hub' (van den Heuvel, 2011) of brain regions "organized in a series of processing areas and nested hierarchies that form networks" linked to extended models of the self (Craig, 2008). For these reasons, as Craig explains, "one might think of the insular cortex as a ventral extension of the sensorimotor strip that is concerned not with proprioception and exteroception, but with interoception" (Craig, 2009).

Interoceptive Dynamics in the Insular Cortex

Originally referred to as the *isle of Reil* (after discovery by German physician and anatomist Johann Reil in 1809), the insula is a hidden cortex located behind the conjunction of the frontal and temporal lobes, in the *fundus* (depth) of the longitudinal (*Sylvian*) fissure. While consisting of finer-scale anatomical details and functional zones, generally speaking the insula can be labeled according to three main regions: *anterior*, *medial* and *posterior*. Sensory-driven interoceptive information in the insula evolves in a *posterior-to-anterior* progression.

Along the spino-thalamo-cortical pathway, afferent C/Aδ-fibers in all bodily organs (etc.) project, in part, to the posterior insula, where modality-specific representations of afferent activity create a temporal representation of objective physical signals from the internal milieu (Craig, 2009). These values proceed through a phase of multimodal 'information integration' with subsequent "inputs from affective, cognitive and reward-related brain circuits" (Seth, 2013), as well as other brain regions — including the *anterior cingulate cortex* (Critchley *et al.,* 2004; Fan *et al.,* 2011; Denny *et al.,* 2012; Lindquist *et al.,* 2012); *amygdala* (Etkin & Wager, 2007); *ventral striatum*; *prefrontal cortex* and primary somatosensory cortex (*S1)* — responsible for sensory information about the external world, bodily representations and past experiences (Mufson & Mesulam, 1982; Craig, 2002, 2003; Harrison *et al.,* 2010). As Arnhart explains:

> The insular cortex receives signals from all the tissues of the body, and these signals are integrated with physical and social stimuli from outside the body and with the memory of past experiences as well as imaginative projections of future experiences. (Arnhart, 2011)

Bud Craig diagrams a heuristic process depicting the general types of information integrated during posterior-to-anterior development of signal processing in the insula.

[3] Damasio uses this term to describe sensory patterns of information—the “most-abundant constituents of mind"

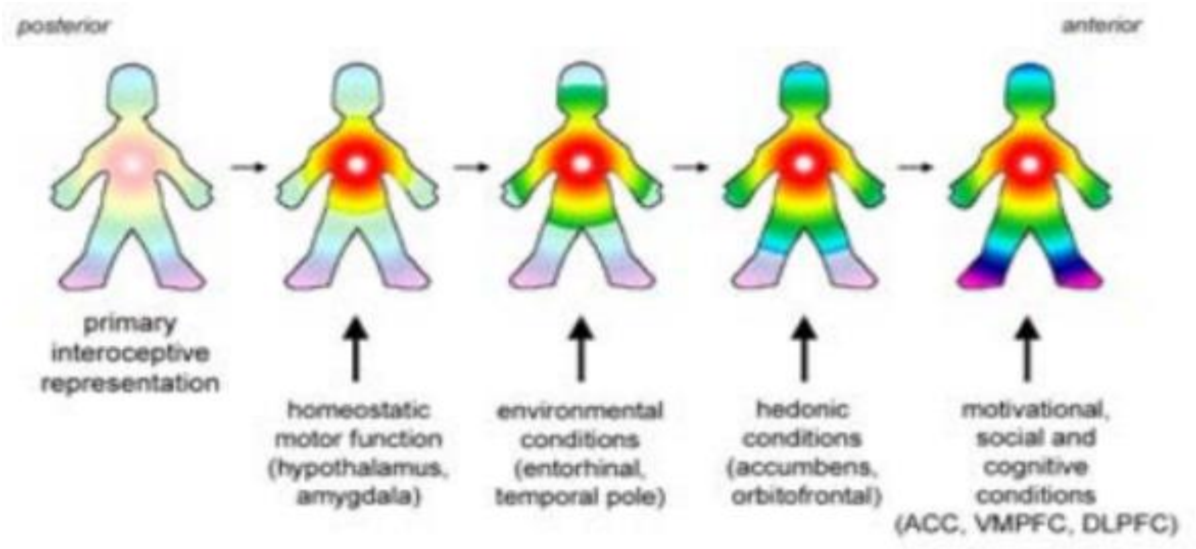


Image reprinted from (Herbert & Pollatos, 2012) as an adaptation of (Craig, 2009). Posterior-to-anterior progression in the insula, with causal and modal correlates

These integrative phases culminate in a "*re*-representation" (Craig, 2009) of multimodal, subjectively-dressed feeling-states in the anterior insula (Parvizi & Damasio, 2001; Craig, 2003; Critchley, 2004; Harrison *et al.,* 2010). Such "sensory patterns of information" (Damasio, 2013) supply not only a cardinal form of self-awareness—as a primary interoceptive image of the internal milieu and "feeling about the homeostatic state of the body" (Seth, 2015)—but also inform other modes of self-awareness like the distinction between self/other implicit in grounding intuitive awareness of complex, fast-paced social interactions (Gu & Hof, 2013; Barlassina & Newen, 2013; Tajadura-Jimenez &Tsakiris, 2014). What is perhaps most noteworthy about this procedure is how associative information is layered-up onto a representational image of raw physical signals (*objective* feelings from the body) to confer a *re*-representation of the objective physical feelings into a *subjective* feeling-state characteristically dressed with an individual's unique experiences.

Interoceptive Inference: Feed-back Processing (Top-Down)

Another program with increasing traction in computational neurobiology is "predictive coding." This top-down model of perception involves the notion of the brain as a "hypothesis generator" (Gregory, 1980; Dayan et al, 2006**;** Knill & Pouget, 2004). Conceived in the West by Rev. Thomas Bayes (Bayes & Price, 1763), and later in the context of vision by Hermann von Helmholtz (1866), the idea is that "perception is inferred unconsciously to correspond to the most-likely physical events or circumstances that could have produced the pattern of sensations that generated it" — called, the "likelihood principle" (Helmholtz, 1866). Such an idea suggests perception requires the management of an internal model. In the present context, by incorporating the "*every good regulator*" hypothesis that to regulate a system requires maintaining a model of that system (Ashby, 1947), it follows that to preserve homeostasis "an organism should maintain well-adapted predictive models of its own physical body (its position, morphology, etc.) and of its internal physiological conditions" (Seth, 2013).

Predictive Processing / Predictive Coding

According to the "predictive coding" model (Dayan *et al.,* 1995; Friston, 2002, 2010), the brain uses internal models to continuously predict sensory inputs "with the imperative to minimize the discrepancy or 'prediction error' between its inputs (*incoming signals*) and its emerging models of the causes of these inputs (*predictions*) via neural computations approximating Bayesian inference" (Seth, 2013; see also Carhart-Harris & Friston, 2010). Bayesian inferences generate predictions (or hypotheses) about the causes of sensations, and in doing so actively generate hypotheses to explain what is happening in external and interior worlds—or to "*explain it away*" (Dennett, 1992). As Barrett and Simmons explain, "predictions function as hypotheses about the world that can be tested against sensory signals that arrive in the brain" (Barrett & Simmons, 2015). The difference between the expectation (prediction) and actual sensory data (e.g., from neurovisceral afferent fibers), is called the "prediction error," which can also be understood as a measure of "free-energy."

Free-Energy Principle & Predictive Coding

Developed by British psychiatrist and computational neuroscientist, Karl Friston, the *free-energy principle* states that any self-organizing system maintaining equilibrium with its environment must also minimize its free-energy (Friston, Kilner & Harrison, 2006). Free-energy is a measure of 'surprise' that correlates with the amount of prediction-error; thus, minimizing prediction-error is effectively the same as minimizing free-energy (Hohwy, 2013). This principle represents a mathematical formulation for how adaptive systems resist a natural tendency to disorder (Friston, 2007; see also Ashby, 1947; Nicolis & Prigogine, 1977; Kaufmann, 1993; Haken & Haken, 2004). While predictive coding models have been applied to *proprioception* and *exteroception*, some consider the paradigmatic application to be realized in terms of *interoception* (Friston & Kiebel, 2009; Gu & FitzGerald, 2014). As Indian-British neuroscientist Anil Seth notes, "one of the most relevant features of the world for a particular organism [to model] is the organism itself" (Seth *et al.,* 2013).

Interoceptive Predictive Coding

Approaching interoceptive dynamics as an "event-logic" set against the backdrop of a self-organizing, regulatory system maintaining an internal environment within fine-tuned parameters, "*events*" may refer to any activity that presents as a *difference* in the internal milieu i.e. a disturbance or departure from the multistable steady-states of homeostatic expectations. Afferent sensory fibers carrying information about the internal conditions of the body (via the *spinothalamocortical* pathway to the posterior insular cortex) are met with converging (top-down) predictions about the causes of those signal-events. Here, events can register as "outliers of interest" or be "explained-away" such that a good deal of innate processing may involve an "explaining away" of mundane events in our environments. Bottom-up "prediction-error" neurons then send the differences back to update intrinsic models.

If the prediction *explains-away* the event, the "prediction-error" is *low* and subsequent neurons project increased confidence and reliability-ratings about the prediction back to intrinsic models (Roitblat, 2013); however, if the prediction is disconfirmed by sensory data (or insufficient to resolve), the prediction-error precipitates an intensification of "salience" corresponding to the variance, and an "alert" signal follows with implicit action policies for attending to the prediction-error. The salience network (*anterior insular cortex*, *anterior cingulate cortex* and *dorsolateral prefrontal cortex*) represents the highest reflex circuit of the neurovisceral axis (Mayer, 2011) that "plays a crucial role in identifying the most biologically and cognitively relevant events for adaptively guiding attention and behavior and constitutes a key interface for cognitive, homeostatic, motivational, and affective systems" (Menon, 2016). As Harrison explains, "consistent with their respective visceromotor and viscerosensory roles apparent across mammalian species," the anterior cingulate and insular cortices are shown to be "reliably engaged during states of enhanced bodily arousal" (Harrison *et al.,* 2010; see also: Critchley *et al.,* 2003; Allman *et al.,* 2005; Butti & Hof, 2010; Evrard, Forro, & Logothetis, 2012). This evolves to enable the orchestrated effect of a constantly adaptive interoceptive sense within a dynamically regulated environment aimed at preserving homeostasis.

This process can also spur subsequent "hypothesis testing" that draws from relevant information across modalities until a new hypothesis (explanation) is derived and adjusted into the intrinsic models—such that models and predictions continuously improve with new data. The process of prediction adjustment is called "precision coding." As Seth explains, "precision weighting (possibly implemented by post-synaptic gain modulation of prediction error units) can modulate the extent to which prediction-errors are resolved" (Seth, 2013).

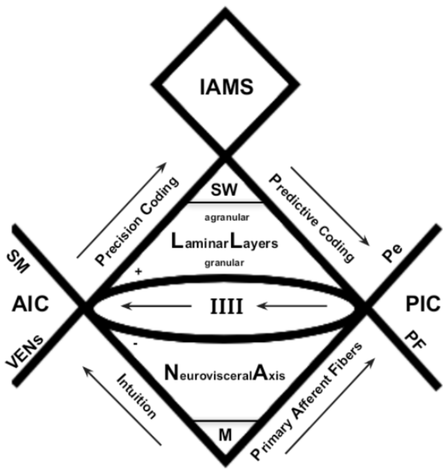
This model provides an example of a natural learning procedure in the "context of discovery" called "*abductive inference*" (Peirce, 1867).

Similar to 'trial-and-error,' this type of inference entails the process of inquiry whereby constraining less-likely options means a shrinking number of more-likely options remains. This leads to better models and inferences. American philosopher C.S. Peirce explains, “abduction is the process of forming explanatory hypotheses. It is the only logical operation which introduces any new idea” (Peirce, 1867). This involves “all the operations by which theories and conceptions are engendered” (*ibid*). This form of inference finds a natural context when cast in terms of the hypothesis-generation role of the brain re: predictive coding.

Conceptual Learning Module

I know it seems like that may have been a lot of content; and, arguably so. However, now I want to show you how to compress the entirety of the previous two sections (and more) into one diagram. The claim is that understanding this diagram will provide a working, conceptual knowledge of interoception. Contextualizing the diagram simultaneously serves as a *formative check* and *conceptual construct* for choosing your own adventure into the materials. Like tasting a fine wine for the first time, as your acquaintance grows, so too you add texture to the taste profile, and gradually come to know and appreciate it in further detail, like tannins of knowledge. This description follows three tracks: bottom-up, top-down and anticipatory (intuitive).

Interoception: Conceptual Platform



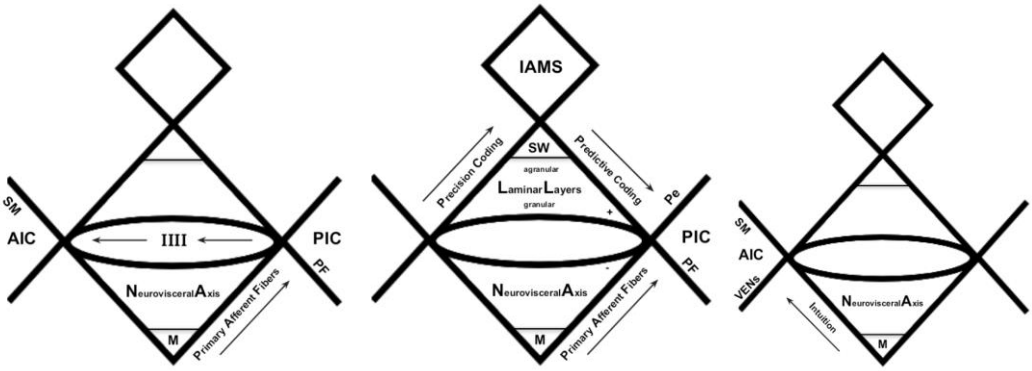
IAMs	-------------- intrinsic adaptive models of self
SW	------------------------------ synaptic weighting
LLL	--------------------------- limbic laminar layers
NV	-------------------------------- neurovisceral axis
PA	--------------------------- primary afferent fibers
PF	---------------------------------- physical feelings
PE	---------------------------- prediction error (+/-)
M	--- microbiome
VENs	------------------------ Von Economo Neurons
SM	---------------------------------- somatic markers
AIC	--------------------------- anterior insular cortex
PIC	-------------------------- posterior insular cortex
IIII	--------- interoceptive info. integration in insula

Starting from the bottom, interoception begins in the context of the *internal milieu*, *microbiome* and regulatory maintenance of *homeostasis*. Within this setup we identify the neurovisceral axis as a ubiquitous information channel specialized to represent one's body from within (Craig, 2002). Activity characterizing the internal milieu involves several capacities (chemical, hormonal, microbial, sensory, cytokine, etc.). Information about this activity transmits to the brain via *spino-thalamo-cortical* paths of primary afferent fibers; e.g., to posterior insular cortex into a representational image of signals as *physical feelings*.

Here there are two options—to continue the bottom-up description or switch to the top-down narrative. I find the former reads easier from a learning perspective. Following the bottom-up account, signals comprising the representational image (in PIC) of *physical feelings* from "the body interior" then transition (in a posterior-to-anterior gradient) through phases of multimodal information integration, where signals are layered-up with associative information from several modalities and then "*re*-represented" in the anterior insula (Craig, 2009). This describes the first track: bottom-up narrative.

The second track again presents a choose your own adventure—you can develop somatic markers and intuition or develop the top-down account. Here I will opt for the latter. To trace-out the top-down narrative begins, fittingly, at the top of the diagram with *intrinsic adaptive models of self* (pronounced like the cat and dog food). The natural imperative for such modeling follows the '*every good regulator*' hypothesis that "every good regulator of a system must also be a model of that system" (Ashby, 1947). In part, these models identify outlier-values from expectations of bottom-up sensory signals from the body (slated to convoke in PIC) — as well as predict subsequent effects of those outliers (said to render in AIC) such as cognitive predictions involving future awareness states (Seth *et al.,* 2011); prediction of risk and error-of-risk relating to unexpected uncertainty (Yu & Dayan, 2005; Alink *et al.,* 2010; Hesselmann *et al.,* 2010). This also adds a pragmatic element to perception in that top-down models only draw from as much sensory data as needed to fill-in the rest.

The difference between the prediction and sensory signals generates a *prediction-error*. If the prediction-error is low, "IAMbic" predictions are updated with an added measure of confidence. If the prediction-error registers high, a salience effect motivates the attention and corrective actions via action policies. Such adjustments are sent back to IAMs in the form of 'precision coding' that updates models via up/down-regulation (gain modulation) of synaptic weights. This is the top-down narrative of track two. The third track involves microbiome communication with the cortex (e.g. protein signatures) underwriting a key part of biological intuition, and with respect to somatic markers, as well as the complementary roles of laminar layers, cortical columns and synaptic weighting, discussed in the further reaches.



Track 1

Bottom-Up

Craig; Mayer

Track 2

Top-Down

Friston; Seth; Barrett

Track 3

Bottom-Up, Anticipatory

Allman; Damasio; Stimpson

Philosophical Foundations of Interoception

In his 1929 book, Process & Reality (PR) Alfred North Whitehead specifies three modes of perception as follows: *causal efficacy*, *presentational immediacy* and *symbolic reference*. Whitehead's 'three modes of perception' are developed in the wider context of a *philosophy of organism* that involves at core a generative process comprising "*actual occasions*." Even without a background in this logic, the following description conveys how interoception (re: neurovisceral axis and insular dynamics) can be shown to apply to the generative logic of *actual occasions*, specifically with respect to "grade-4" of Cobb and Griffin's model (1976). In this section, to observe due diligence, references to (Whitehead, 1929) are specified by page number as follows, (PR, nn).

> The occasions of experience are of four grades [...] Occasions of experience of the fourth grade involve experience in the mode of presentational immediacy, which means more or less what are often called the qualia of subjective experience. So far as we know, experience in the mode of presentational immediacy occurs in only more evolved animals.

Reading "presentational immediacy" as "interoceptive awareness" (of the *milieu intérieur*, i.e. Whitehead's mode of "causal efficacy") confirms the evolutionary adaptation of this form of awareness (interoception as the presentational immediacy of the internal milieu) as we are reminded from Craig's quote (in the introduction) regarding the neurovisceral axis as "a phylogenetically-novel homeostatic sensory afferent pathway in primates, especially [hominids] that provides the basis for a sense of the physiological condition of the body" (Craig, 2002) entailing a more-evolved subjective experience (qualia).

For a quick note: as we have learned so far, interoceptive awareness involves converging operations of top-down and bottom-up signal processing (Barrett & Simmons, 2013). When thinking about these causal processes in terms of Whitehead's program, there is a key difference to address; namely, where predictive coding operates on the basic imperative to minimize surprise (explain features away)—Whitehead's model, re: experience-as-lived, corresponds mostly with bottom-up causality where complexity correlates with the vividness of experience. These two imperatives are distinct, and yet also complementary in that, as Friston explains, minimization of free-energy also corresponds with a maximization of evidence (Friston, 2010). In this sense, the latter also resembles a measure of consciousness by what Tononi defines with the scalar operator (Φ) to denote the degree of information-complexity or richness of subjective experience.

Causal Efficacy (CE)

Causal efficacy refers to "the bodily feeling we have of being affected-by and dependent-on an ambient physical milieu, as well as the feeling we have of the organic body itself as the most proximate milieu of this sort" (Weber & Weekes, 2009). As Whitehead states, "bodily experiences, in the mode of causal efficacy, are distinguished by their comparative accuracy of spatial definition" (PR, 176). Such a milieu clearly refers to Claude Bernard's "*milieu intérieur*" (Bernard, 1865) operating within the regulatory oversight of what Walter Cannon coined *homeostasis* (Cannon, 1929). Similarly, as Whitehead explains, "the original locus of perception is not in consciousness but in the organism's manifold forms of autopoiesis insofar as this autopoiesis is necessarily congruent with the environment in which it manages to survive" (Weber & Weekes, 2009). This describes what can be understood through the lens of variational homeostatic regulation (or allostasis) vis-à-vis anticipatory monitoring and adjustments of the internal milieu that "facilitate[s] compensatory changes supportive of physical and physiological functioning essential to survival" (Turrigiano, 1999). Weber and Weekes summarize this in the following quote;

> The organism adapts itself moment by moment to ambient influences from its immediate environment. This happens in internal processes of homeostasis that continually compensate for fluctuations in critical parameters of the organism's milieu, as well as in reflex movement and the many forms of arousal and affectivity that are elicited biochemically and do not depend on conscious mediation. In all of these self-adjustments the organism is representing its environment (Weber & Weekes, 2009).

Weekes describes how perception in the mode of *causal efficacy* is "characteristically vector-like in conveying vague but imperative information [...] impinging on the present actuality of experience" (Weber & Weekes, 2009). Another way to say this is that the information follows a certain pathway-like vector. Weekes suggests this 'vectorial activity' is "unspecific and diffuse, but still veridical." In a neurovisceral context, we can qualify these terms ("*unspecific,*" "*diffuse*") in terms of the ubiquitous multiscale dynamics of afferent sensory fibers enervated in every organ at different nexüs of nodes tracing two primary pathways to the cortex (the "*veridical*" aspect); therefore, while diffuse throughout the organs and tissues of the body, the spinothalamocortical progression of afferent fibers to specific targets in the cortex describes a vectorial (pathway) component. Without identifying the neural modality, Weber and Weekes are keen to recognize the contents as "mostly visceral in locus and affective in content" (Weber & Weekes, 2009). They explain, a "physical process at its most elementary level is a faint wisp of affective experience 'enjoyed' by an extremely rudimentary nonconscious subjectivity" (*ibid*). This type of construct is demonstrably covalent with activity conveyed by primary afferent fibers (etc.) found in every organ and tissue of the body comprising the dynamic luminal context of the body's *interior milieu*—and comparable to the closely-related neurophenomenological view claiming "affect and emotions as the originary source of the living present" (Depraz & Varela, 2005); and to the description of "primitive self-consciousness [as] fundamentally linked to bodily processes of life-regulation, emotion and affect, such that all cognition and intentional action are emotive" (Lutz & Thompson, 2003; see also Marstaller, 2009).

Presentational Immediacy (PI)

Perception in the mode of "presentational immediacy" involves "a world decorated by sense-data dependent on the immediate states of relevant parts of our own bodies" (Whitehead, 1927) re: *internal milieu* (mode of causal efficacy). Specifically, "presentational immediacy is to be conceived as originated in a late phase, by the synthesis of the feeling of bodily efficacy with other feelings" (PR, 312). "This 'with-ness' of the body is an ever-present, though elusive, element in our perceptions of presentational immediacy. The feeling of bodily efficacy [...] is more primitive than the feeling of presentational immediacy which issues from it. Both in common sense and in physiological theory, this bodily efficacy is a component presupposed by the presentational immediacy and leading up to it" (PR, 312). As such, "presentational immediacy begins with sense presentation of the contemporary body" (PR, 81) where "our bodily experience is primarily an experience of the dependence of presentational immediacy upon causal efficacy" (PR, 176).[4] To validate this claim, Whitehead explains

[4] "Hume's doctrine inverts this relationship by making causal efficacy, as an experience, dependent upon presentational immediacy. This doctrine, whatever be its merits, is not based upon any appeal to experience" (176).

how “the predominant basis of perception is perception of the various bodily organs, as passing on their experiences by channels of transmission and of enhancement” (PR, 119). Here, "the present perception is strictly inherited from the antecedent bodily functioning, unless all physiological teaching is to be abandoned" (PR, 312) meaning that "the perceptive mode of immediate presentation affords information about the percepta in the more aboriginal mode of causal efficacy" (PR, 178).

This occurs “by means of our projections of our immediate sensations, determining for us characteristics of contemporary physical entities” (Whitehead, 1927). As Whitehead describes, “presentational immediacy is an outgrowth from the complex datum implanted by causal efficacy” (PR, 173). “The function of the phase of presentational immediacy is to provide a representation or mapping of the datum, which is dimly discerned at the level of causal efficacy” (Maclachlan, 1992). In this sense, *presentational immediacy* describes the ongoing representational images of the internal milieu expressed as the "*complex datum*" per the luminal basis of perception of one’s interior milieu as a complex of physical feelings re: sensory fibers represented in the posterior insula. "The percepta in the mode of presentational immediacy [...] in comparison, are distinct, definite, controllable, apt for immediate enjoyment, and with the minimum of reference to past, or to future. We are subject to our percepta in the mode of efficacy, we adjust our percepta in the mode of immediacy" (PR, 179).

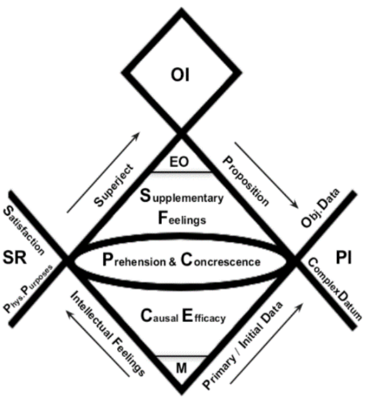
Symbolic Reference (SR)

“When human experience is in question, ‘perception’ almost always means perception in the mixed mode of symbolic reference” (PR, 168). Specifically, "symbolic reference is the interpretative element in human experience" (PR, 173). This “requires a common ground connecting the two pure modes of experience" (Maclachlan, 1992). As Whitehead states, "by this necessity for a 'common ground' it is meant that there must be components in experience which are directly recognized as identical in each of the pure perceptive modes" (PR, 168). This refers to a synthesis of the modes of *causal efficacy* and *presentational immediacy*. As Whitehead explains, "symbolic reference, though in complex human experience it works both ways, is chiefly to be thought of as the elucidation of percepta in the mode of causal efficacy by the fluctuating intervention of percepta in the mode of presentational immediacy" (PR, 178). "This fact, that 'presentational immediacy' deals with the same datum as does 'causal efficacy,' gives the ultimate reason why there is a common 'ground' for 'symbolic reference.' The two modes express the same datum under different proportions of relevance" (PR, 173).[5] Applied in terms of neuroscience, "feelings, and thus consciousness, arise through a comingling of signals from the interior of the body with signals from the nervous system. This is a critical idea because most accounts of consciousness rely exclusively on neural processing rather than an interaction between neural and body processes" (Damasio, 2021).

In terms of the neurovisceral axis underwriting interoception, afferent signals originating from all body organs of the internal milieu (Sterling & Eyer, 1988) project to PIC to take on other values before re-representation in AIC as subjective "feeling-states" of awareness (Craig, 2009). In this context, *symbolic reference* could refer to the symbolic *re*-representations (maps, images) of the state of the subjective self, or what we might call a *subjective symbolic reference*. What characterizes these "re-representational images" (*ibid*) is the patent subjectification of objective signals from the body, layered-up (or "*concrescing*") via a plethora of *associative*, *cognitive*, *affective* data ("*supplementary-, intellectual-, primary- prehensions*") from past experiences (memory, models) plus *intuitions*. Whitehead's three modes parameterize general characteristics of the neurovisceral axis and insula in terms of interoception so precisely they can be regarded as formal philosophical foundations.

Whitehead's ideas in the Model

Beyond a demonstrable overlap in language and ideas befitting a level of recognition as formal philosophical foundations for interoception, another layer of support and connections can be realized by applying the term-logic of AE's to the learning model.



CE -------------------------------- mode of *causal efficacy*

PI -------------------- mode of *presentational immediacy*

SR --------------------------- mode of *symbolic reference*

P&C ---------- "*prehension*" & "phases of *concrescence*"

OI -------------------------------- "objective immortality"

PhyPur ------------------------------ "physical purposes"

Whitehead's "*initial data*" to "*objective data*" transform ≡ afferent fibers in neurovisceral axis to posterior insula

[5] "So far as concerns conscious judgment, the symbolic reference is the acceptance of the evidence of percepta, in the mode of immediacy, as evidence for the localization and discrimination of vague percepta in the mode of efficacy" (PR, 179).

Initial data ≡ mode of *causal efficacy*

Objective data ≡ *presentational immediacy; pret-a-prehension*

Prehension and *concrescence* take place in IIII development of signals from PIC to AIC

Information reaches "*satisfaction*" as subjective, symbolic re-representational references to objective bodily feelings.

One way to develop a narrative from this diagram, as described in the '4th-grade of AO's' by Cobb & Griffin, is to begin with Whitehead's "mode of *causal efficacy*" re: neurovisceral axis, homeostasis of the internal milieu, and interoception. In this context, cortically-projecting (*afferent*) sensory fibers may correlate with what Whitehead refers to as a "*multiplicity of initial data*." This generates a "*complex datum*" as a physical image (or sensory-mapping) of the objective internal milieu of the body in the posterior insular cortex, linked to perception in the mode of *presentational immediacy*. Following the model it would appear there are two different features we can describe in the mode of presentational immediacy. This is validated in Whitehead (PR, 173) where he explains how:

> The two genetic processes involving presentational immediacy must be carefully distinguished. There is first the complex genetic process in which presentational immediacy originates. This process extends downwards even to occasions which belong to the historic routes of certain types of inorganic enduring objects, namely, to those enduring objects whose aggregates form the subject-matter of the science of Newtonian dynamics. Secondly, prehensions in the mode of presentational immediacy are involved as components in integration with other prehensions which are usually, though not always, in other modes. These integrations often involve various types of 'symbolic reference.'

In Whitehead's account, *initial data* "transition" to "*objective data*" available for the phases of "prehension" and "concrescence" (Lango, 1972), as a selective integration of associative information. As such, "*objective data*" may refer to primary afferent fibers in the posterior insular cortex as they begin the posterior-to-anterior progression-cycle of multimodal information integration (phases of "prehension" and "concrescence"). Prehension involves "primary" feelings (affective); "supplementary" feelings (associative); "comparative" feelings—of which there are two types, one "intellectual' feelings (cognitive) subdivided into two species: one consisting of '*conscious perceptions*'; and the other of '*intuitive judgments*' (PR, 266); and "the other type termed '*physical purposes*' that arise from the integration of a conceptual feeling with the basic physical feeling from which it is derived" (PR, 266). Each of these has a correlation to be developed in terms of interoception. This "prehensive" integration proceeds in progressive "concrescence" until culminating in a "*re-representational* *image* in AIC" (Craig, 2009), as a "symbolic reference" to the gestalt of subjective feeling-states (what Whitehead calls a "satisfaction") involving and evolving in the context of interoceptive awareness, subjectivity and emotional feeling-states. The "*superject of satisfaction*" (Whitehead) fits nicely with the idea of how the "*satisfaction*" variously impacts future states based on *precision coding* adjustments to the intrinsic models of each individual during the *superjective* phase. In many ways, these intrinsic models supply the most "objective" and "immortal" information about each person.

Corollary features appear in the top-down model of interoception. For instance, Whitehead describes a "propositional feeling" (PR, 266) in terms that comply with a Bayesian model. As he states, "each successive propositional phase is a lure to the creation of feelings which promote its realization" (PR, 225)—i.e. predictions (*propositions*) about the causes (or *creation*) of sensory signals entail those sensory signals they expect to see. Such predictions emanating from deeper laminar layers may also conform with Whitehead's "*eternal objects*" as synaptic weights linked to "objective immortality" (intrinsic models). By contrast, values in the superficial layers of the lamina correspond with the granule cells of sensory data likened to the mode of *presentational immediacy*. The beauty of this approach is that technical terms in neuroscience can be superimposed onto terms in Whitehead's philosophy such that you can learn about one from knowing about the other.

Whitehead's Terms	Neuroscience, Interoception
• Caual Efficacy -------------------------	• Neurovisceral Axis, Honeomsotasis of Internal Mlilieu
• Multiplicity of Initial Data -------------------	• Primary Afferent fibers → PIC (spinothalamocortical)
• Presentational Immediacy ------------------	• Representational image of physical signals from body
• Objective Data ------------------------	• Primary C and Aδ afferent ensory fibers in PIC
• Prehension & Concrescence ---------------	• Multimodal information integration in the insula
• Complex Datum ------------------------	• Representational image of C/Aδ afferent fibers in PIC
• Physical Feelings -----------------------	• Comprising C/Aδ afferent sensory fibers in PIC
• Supplementary Feelings -------------------	• Associative information layered onto 'physical feelings'
• Intuitive Judgments ----------------------	• Intuition conferrred as gut feelings in neurovisceral axis
• Satisfaction -------------------------	• Subjective re-representation of physical feelings in AIC
• Symbolic Reference ---------------------	• Subjective feeling states symbolic of physical feelings
• Propositions ---------------------------	• Generative "IAMbic" predictions and expectations
• Superject of Satisfaction ------------------	• Precision Coding
• Objective Immortality -------------------	• Intrinsic Adaptive models of the self (Iams)
• Eternal Objects ----------------------	• Synaptic Weighted probbilities evolving in time

Discussion

While I draw heavily from other's work—in particular, Craig, Friston, Weekes, Seth, Barrett, Allman and Whitehead—overall this speaks to the ability to bring together the research trajectories of several scholars across fields and disciplines in their own words, under this topic. In many ways I am just putting the various pieces together and pointing to connections that others can draw-out and fine-tune in a more-formal capacity. Still, this review is not without original contributions. In addition to the conceptual learning model, a showcase (or testing-ground) for Tononi's "Integrated Information

Theory" of consciousness (IIT) may be identified in terms of a model of Interoceptive Information Integration in the Insula (IIIIT).

Interoceptive Information Integration in the Insula

One of three leading candidates for a physical theory of consciousness [that since first writing this article (in 2022) has recently come under pressure in a provocative open letter], the "integrated information theory" of Guilio Tononi seeks to explain how subjective experience, or "how it feels to be you here and now in each moment," can be put into physical terms and localized in neural operations (Tononi, 2021).[6] To model this he constructs "irreducible complexes"[7] as mathematical correlates of subjective experience whose information involves a matter of *content*: what it is; *qualia*: the particular feelings it evokes; and *mechanism*: to what extent various parts of the brain (and intrinsic connectivity networks) represent activity and values correlated with subjective experience (Tononi, 2019).[8]

When reframed in terms of express goals, IIT and IIIIT (pron. "yeet") both seek to account for the phenomenology of subjective experience (qualia). In this way, Tononi's two goals match with the directives of Bud Craig and Emeran Mayer (in terms of mapping neurobiological mechanisms and dynamics of the neurovisceral axis) to address *subjective experience*—namely, "how do I feel now?" (Craig, 2003; 2009). In particular, both approaches seek to describe the ongoing expression of subjective feeling-states of experience (comprising "what it is like to be you"). While the methods are different, Tonini's goals are shared. Plus, both trajectories can be linked to a shared topic: *subjective experience*, and to a shared feature: *representational constructs* of subjective experience.

In terms of interoceptive information integration in the insula, multimodal inputs are integrated with objective physical signals from the body to "layer-up" a subjective feeling state and enhanced sense of acquaintance and ownership. In the same way Tononi identifies the richness of an experience as proportional to the complexity of

[6] The theory of integrated information has gone through four major iterations. The idea of complexity as a measure of shared information began with Tononi and Edelman in 1998. Then, in 2003 Tononi put forward a model of IIT in terms of "effective causal information" of the whole over the parts, in a stationary model of integration. Five years after that he updated the model to a dynamical setting (Rosas & Chandaria et al., 2020).

[7] IIT 4.0 introduces *partition functions* quantifying integrated information by the sign *phi* (Φ) equal "to what extent a system is physically irreducible" in terms of connections that, if broken, make a difference in state. This measure addresses what the system says about itself and how much. To model this, Tononi constructs a "maximally irreducible conceptual structure" as a triadic relation that models a "complex" or "quale" (Tononi, 2019).

[8] Tononi names *five* axioms of consciousness as "essential properties that every experience must satisfy," akin to Whitehead's *speculative philosophy*. These are: 1. Experience exists. 2. Experience is structured; it has composition and aspects. 3. Experience is differentiated; it is particular. This defines *information*. 4. Experience is unified; presented as a gestalt that has *causal efficacy*. This defines *integration*. 5. Experience involves an aspect of *exclusion*, namely, that there are non-arbitrary constraints placed on states so that not all possible states apply.

'*integrated information*' (Φ), so too we might imagine a corollary richness of subjective experiences linked to the degree of integrated information layered in the anterior insula. Information integration in AIC is supported by anatomical (Mesulam & Mufson, 1982b; Saper, 2002) and intrinsic functional connectivity (Cauda *et al.,* 2011; Deen *et al.,* 2011; Touroutoglou *et al.,* 2012; Chang *et al.,* 2013) with a large-scale network of sensorimotor, affective, and cognitive regions. The signal diversity in AIC explains the involvement across low-level *autonomic* and *sensorimotor* aspects (Craig *et al.,* 2000; Sterzer & Kleinschmidt, 2010; Fan *et al.,* 2011) to high-level *cognitive* and *social* domains (Montague & Lohrenz, 2007; Rilling & King-Casas *et al.,* 2008; Bossaerts, 2010; Kirk *et al.,* 2011). In addition, as Mayer explains, "recent evidence suggests that various forms of subliminal interoceptive inputs from the gut, including those generated by intestinal microbes, may influence memory formation, emotional arousal and affective behaviours. The human insula, and related brain networks (including the anterior cingulate cortex, orbitofrontal cortex and amygdala), has emerged as the most plausible brain region to support this integration" (Mayer, 2011).

Tononi and others have in mind structures in the brain that demonstrate "causal emergence" (Rosas & Chandaria *et al.,* 2020). The neurovisceral axis, and specifically the integration of interceptive information in the insula, offers a sense of *causal emergence* or *efficacy* insofar as physical signals from the body are layered-up into subjective proprietorship. What emerges are feelings of subjectivity, ownership, etc. This provides a biological example of the computational model of Tononi's "integrated information" in terms of "interoceptive information integration in the insula," where both approaches seek to describe the ongoing generation of unique qualia characterizing subjective experience.

Given Tononi's two objectives for IIT (to be put into physical terms and localized in neural operations), IIIIT would appear to satisfy both. Still, what makes for good theory and good science is testability (Popper, 1963). If *subjectivity* is an "essential axiom… something that needs to be explained" (Tononi, 2019), given Tononi's approach begins from phenomenology (*experience itself*, Rosas & Chandaria, 2020) and his category of "*mechanism*" seeks neural correlates, it may be plausible to test his computational procedure of "qualia structuralism" to see if it is possible to model the qualitative complexity (or richness) of subjectively-dressed interoceptive feelings in an "isomorphic" mathematical representation space of the insula (Oizumi & Tononi, 2014; Krauss & Maier, 2020).

Further Reaches

The remaining track in the conceptual diagram of interoception includes the corollary features of *somatic markers* in AIC; *intuition* in the context of the microbiome and neurovisceral axis; specialized *Von Economo neurons* in AIC and ACC linked to the gut and intuition; and *laminar layers of limbic cortices*. An overview of each is provided here.

Somatic Markers

Predating Craig's idea of "interoceptive images" (Craig, 2003) and packet-like events of “*emotional moments*” constantly expressed in converging multimodal nodes of network activity in the insula (Craig, 2011), in 1994 the neurobiologist Antonio Damasio introduced a “somatic marker hypothesis” (Damasio, 1994; 1996, 2009) in the context of feelings, emotions and decision-making. Damasio emphasizes that "interoception, which means bodily responses such as sympathetic activity, can be represented in the insula and anterior cingulate cortex and can play critical roles in decision-making" (Ohira, 2015). According to this theory, "automated signal" associations of past memories decorate "raw physiological signals" (*somatic markers*) from the body that "influence emotional states and bias our inclinations when making decisions" (Bechara, 2004; see also Bechara, A.Damasio & H.Damasio, 2000; Prinz, 2004; Dunn, Dalgeish & Lawrence, 2006).

These "*in-tarot-ceptive*" markers facilitate adaptive cognition and emotional behaviors in social situations, "increas[ing] the accuracy and efficiency of the decision-making process...and in their absence diminishes it" (Damasio, 2021). The idea is *somatic markers* aid cognitive processes that alone could not handle the computational load required for adaptive responses. In this way, somatic markers can be thought of as an extension of interoception associated with gut feelings from the viscera that provide useful markers for interpretation. The interplay between individual differences in the generation of somatic markers and the ability to perceive subtle bodily changes of an interoceptive nature can also account for significant variability in human intuition (Dunn & Dalgleish *et al.,* 2010).

Intuition and Gut Feelings

The biological basis for intuitive decision-making involves a cortical representation of affective markers predicated on visceral activity[9] in

[9] The human body is like a super-complex ecosystem containing trillions of bacteria and other microorganisms that inhabit all our surfaces including the skin, mouth, sexual organs, and especially the epithelial walls of our intestines (Pacheco-López, 2013). In extreme estimates, microbes outnumber cells on the order of ten-to-one (Wenner, 2007), suggesting we may be more microbial than cellular. This suggests an important role for the enteric nervous system and microbiome, and a corresponding impact from molecules and neurotransmitters from the gut such as serotonin that exhibit mood effects based on luminal conditions. Given rudiments of the enteric nervous system are found throughout the animal kingdom, including in insects and

concert with cellular adaptations enabling fast-paced decisions from interoceptive stimuli (Allman *et al.,* 2006; Preuschoff *et al.,* 2008). As Dunn explains, "emotional experience and intuition are associated with individual differences in the ability both to generate and to perceive accurately subtle changes in the body" (Dunn, 2010).[10] As Caltech neuroanatomist John Allman explains:

> Intuition is a form of cognition in which many variables are rapidly evaluated in parallel and compressed into a single dimension. This compression facilitates fast decision-making. Typically we are not aware of the logical steps or assumptions underlying this process although intuition is based on experience-dependent probabilistic models. Instead we experience the intuitive process viscerally. Intuition operates largely in the social domain but can also be applied to purely physical situations. Intuition is plastic; it is not instinct, although instinctive feelings may contribute to it. Emotional value judgments contribute to both intuition and deliberation (Allman *et al.,* 2005).

As we learn to match cognitive (more generally, *existential*) information with immediate and subtle, context-based affective signals, so too do we develop our intuitive repertoire. Wisdom, Damasio says, is cultivated in large part alongside the management of these inner emotions (Damasio, 2021). There is evidence this also translates into observable variations in neural-type and volumetric density in key brain regions.

Von Economo Neurons

Cytological evidence spearheaded in the early 2000's by John Allman suggests that evolutionarily-unique *Von Economo* neurons (VENs) appearing most-abundantly in human interoceptive cortices are key to enabling the task-demands of complex, fast-paced processing needed for rapid decision-making linked to social dynamics and intuition.

Found only in a handful of species with highly-complex social structures, brain size and encephalization ratio (brain-to-body size), VENs are *evolutionarily*, *morphologically*, *regionally*, *functionally* and

snails (Mayer 2011), it is suggested the ganglia that eventually formed the higher cortices originated in more-primitive enteric circuits (Bullock & Horridge, 1965).

[10] For instance, cross-frequency electrical dynamics of neural oscillations are purported to play a functional role between computational communications of large-scale brain activity at the behavioral level, and local cortical processing at the synaptic level (Canolty et al, 2012; Aru et al., 2015; Yufik, Sengupta & Friston, 2017). To these ends, a recent experiment by (Richter et al., 2016) finds clear evidence of a *correlation* (if not 'coupling') between the ongoing gastric basal rhythm of the stomach, and alpha frequency values (10-11 hz) of brain wave oscillations in the right anterior insula. This result extends the notion of *phase-amplitude coupling* to the neurovisceral axis and further suggests that the gastric basal rhythm, which is driven by primary afferent fibers ascending the neurovisceral axis, provides another key layer in mood and emotional processing.

immunohistochemically specialized neurons appearing late in ontogeny and phylogeny—around 15,000 and 30,000 years ago in a convergent evolution of pachyderms/cetaceans and great-apes/hominids (Allman et al, 2002) and most-expressed in the human brain (Allman, Watson & Tetreault, 2005; Watson *et al.,* 2007).[11] Likely a morphological specialization of large pyramidal cells, VENs possess longer axons for faster signal transmission, and where the majority of growth is concentrated in a single, massive dendrite (similar to a motor projection neuron). As Allman explains, “this specialization enables us to reduce complex social and cultural dimensions of decision-making into a single dimension that facilitates the rapid execution of decisions” (Allman, 2005b).

Regionally, VENs are only found in three areas: *anterior insular cortex*, *anterior cingulate cortex* (ACC), and *dorsolateral prefrontal cortex* (dlPFC). Within these regions, VENs only appear in laminar layer 5b—an output layer—thereby reinforcing the idea these neurons are topologically positioned to receive an array of diverse information and compress it for rapid and diffuse relay to other regions by a simple-signal (Allman *et al.,* 2005). In terms of molecular profile, VENs express antigen receptors associated with the transmission of signals from the gut and social bonding (Allman et al, 2011; Watson *et al.,* 2006; Allman & Watson, 2007).[12] Notably, AIC and ACC "help to create a multisensory representation of the body in the world so that what you see, hear and feel is influenced by your interoceptive predictions" (Barrett & Simmons, 2015). This supports the swift transmission of visceral afferents to the brain, and diffusely throughout the brain, to navigate fast-paced and complex dynamics characteristic of survival and social interactions that the gut can process on a level superior to cognition (Allman, 2011; Kuo *et al.,* 2009; Mayer, 2011). As Allman explains:

> Von Economo neurons may relay a vast intuitive assessment of complex social situations to facilitate the rapid adjustment of behavior in quickly changing social situations. Thus they may be a neuronal specialization linked to the increased complexity of hominoid and especially human social interactions. We hypothesize that this specialization enables us to reduce complex social and cultural dimensions of decision-making into a single dimension that facilitates the rapid execution of decisions. Other animals are not encumbered by such elaborate social and cultural contingencies to their decision making and thus do not require such a system for rapid intuitive choice.

[11] ~16,500 in gorillas to over 193,000 on average in adult human brains (Allman, 2005).

[12] In particular, dopamine 3, a high-affinity receptor expressed *only* on VENs, plus Serotonin 2b. These receptors are rarely found in the central nervous system, instead in the mesenteric (peripheral) nervous system where they regulate gastric motility and visceral feelings. VENs also express Neuromedin B, involved in gastric motility. The expression of all three implicates a relationship with the gut. A fourth neurotransmitter strongly expressed on VENs is the vasopressin 1a receptor, corresponding with mediating the formation of social bonds (Stimpson et al., 2011).

> Because of their late emergence in phylogeny and ontogeny, the von Economo neurons may be particularly vulnerable to dysfunction. (Allman *et al.,* 2005)

Involvement of VENs in the diffuse projection of information throughout the brain also complements Tononi's view that "consciousness is associated with a widespread correlation of information...where information of the parts is shared across the whole" (Rosas & Chandaria *et al.,* 2020). Another contributing factor may entail how "the electrodynamics of the conscious brain are poised near a critical phase transition "and that this near-critical behavior supports the vast flow of information through the brain during conscious states" (Toker *et al.,* 2022; see also Deco & Jirsa, 2012).

Limbic Laminar Layers, Interoceptive Predictive Coding & Synaptic Weighting

As a final topic to mention here, localizing interoceptive dynamics in the insula is a sort of toy model. Barrett and Simmons suppose that an operation characterizing so much about subjective experience seems evolutionarily unlikely to occur in just one brain region. Rather, they extend the idea of predictions issuing from *agranular cells* in deeper laminar layers of the insula (Friston, 2010) to include agranular visceromotor cortices across all limbic regions that "send predictions to and receive prediction-error signals from cortices with greater laminar differentiation in an effort to create the kind of synchronized brain activity that is necessary for consciousness" (Barrett & Simmons, 2015).[13]

For background, the basic functional unit of the cerebral cortex is structured according to vertical columns[14] with different numbers of horizontal layers or *laminae* (Beul & Hilgetag, 2015), where each layer has "characteristic cell types and patterns of intracortical and intercortical connectivity" (Barrett & Simmons, 2015).[15] Sensory fibers register on supragranular layers of the lamina; predictions derive from deeper, agranular cell layers.

[13] "As predictions propagate across cortical regions...they modulate the firing of neurons within cortical columns in anticipation of these regions receiving actual sensory input from the environment" (Barrett & Simmons, 2015). Predictions converging on sensory signals send back a prediction-error, therein inciting differential activity in the anterior cingulate cortex in order to combine to form a salience signal, as well as updating models reflected in synaptic weighting "to sculpt future outputs and predictions" (Barrett & Simmons, 2015).

[14] "Vernon Mountcastle was the first to propose that a canonical circuit consisting of cortical columns underlies everything the neocortex does. The way we see, feel, hear, move, and even do high level planning runs on the same circuitry. If we can understand how a single cortical column works, we will have a framework for understanding how the entire neocortex works." (Hawkins et al., 2018)

[15] Upper three layers (closer to the cortical surface) are called the supragranular layers (I-III) and are functionally distinct from the lower two layers (V-VI), called the infragranular layers (Wurtz & Albano, 1980; May, 2006).

In Lisa Feldman Barrett's "EPIC" model, *predictions* and *prediction-errors* flow within the laminar architecture of connections between cortices where "agranular visceromotor cortices contribute to interoception by issuing interoceptive predictions" (Barrett, 2015). This sequences a logic whereby agranular cells of deeper laminar layers project predictions onto highly-granulated sensory cells in PIC, thus providing a communication gradient in laminar cytoarchitecture from less-laminated to more laminated areas (Friston & Kiebel, 2009; Chanes & Barrett, 2016). "Precision cells tune the gain on predictions and prediction-error dynamically, thereby giving these signals reduced (or, in some cases, greater) weight depending on the relative confidence in the descending predictions or the reliability of incoming sensory signals" (Barrett & Simmons, 2015; see also Seth, 2013).

As Barrett and Simmons explain, the entire limbic belt is comprised of agranular cortices, and this speaks to a certain phenomenology associated with "images" and models (2015). Given the rich-club hub (van den Heuvel & Sporns, 2011) of AIC, ACC, etc., "the interoceptive system provides efference copies to multiple sensory systems and thus forms the basis for unified conscious experience" (Barrett & Simmons, 2015). In addition, such a "noisy signaling environment [is] ideal for transient synaptic weighting to provide the ability to recall global mappings and concepts (neural groups) relevant to ongoing processing" (Edelman, 2006; Freeman & Vitiello, 2006; 2008). According to this approach, "what you experience is in large part a reflection of predictions from agranular limbic cortices (Barrett & Simmons, 2015). Together, these four elements constitute the third track.

Interoception in the range of 'primary consciousness' (Edelman, 2004), the 'proto-self' (Damasio, 2003); and subjective awareness of 'the material me' (Craig, 2003) refers to an ability to sense the changing interior physiological conditions reflected in the homeostatic monitoring of the body. These ever-changing internal conditions can be referred to as the *internal milieu* (Bernard, 1865). As Craig explains, "the neural substrates responsible for subjective awareness of one's emotional state are based on cortical representations of their physiological state" (Craig, 2003) that can be contextualized as the result of a multimodal integration process of associative information in the insula (Craig, 2002; Arnhart, 2011). This can be evinced in terms of language as well.

> When we use the language of physical pain to describe our social pain ("a broken heart"), we recognize the embodiment of our natural social consciousness, in which our mind, our brain, our body, and our social life are inseparably intertwined. (Arnhart, 2011)

A review follows in the form of responses to the four purposes of this paper [*see intro*].

A. While bottom-up models explain perception classically in terms of feed-forward causal feature detection, top-down methods posit inference-generating models operating in a predictive capacity on the causes of incoming sensory signals (Friston, 2012; Picard & Friston, 2014). These two aspects of interoception play integral and complementary roles. In the bottom-up model, afferent sensory fibers in the neurovisceral axis conveying feeling-states about the body's internal milieu ascend in reflex loops from the microbiome and enteric nervous system to a cortical representation in the posterior insula. From here, objective physical signals are progressively layered into subjective *re-representational images* and feeling-states in the anterior insula. This multimodal information integration furnishes subjective experiences in each moment with cognitive and qualitative features.

Another camp of scholars maintains that predictions about sensory inputs generated from intrinsic models are responsible for filling-in and guiding a large part of the contents of experience and perceptions based on inferences to the best guess (Clark, 2013). As Friston explains, "*feedforward architectures on their own are not sufficient [...] feedback connections mediating internal or generative models of how sensory inputs are caused are essential*" (Friston, 2002). In this model, primary afferent fibers projecting to the posterior insula are assessed against inferential hypotheses of what the brain predicts is going on inside the body, based on expectations of past experiences (Seth, 2013). This updates classical models of interoception to state that "interoceptive perception is largely a construction of beliefs that are kept in check by the actual state of the body" (Barrett & Simmons, 2015). Here, the main goal is to control the system's essential variables. Interaction with the world is only necessary to the extent that it affects these variables (Wiese, 2014).

To provide a covering logic for this relationship, we turn to the 20th-century American author, essayist and poet, Robert Frost, who offered the following response when asked by a reporter (at a press-junket on the eve of his eightieth birthday), "what is freedom?" — to which Frost replied, "*It is running easy in your harness*" (1954). Such an idea may also inform the present context. In the bottom-up dynamics, "*running easy in your harness*" may refer to internal operations maintained within optimal survival parameters, while in top-down dynamics, "*running easy in your harness*" may refer to the inferential predictions of generative models largely matching in expectations with the causes of incoming sensory signals, and thus minimizing free-energy. Indeed, the easier and more-freely we 'run in our harness,' so too the more we are able to stabilize and expand the physiological parameters required for the flourishing of more-sophisticated emotions.

B. Accompanying the experiential faculty of interoception is a storehouse of technical topics serving as a conceptual locus for learning and comprehension. The accompanying conceptual model provides an advanced learning device remarkably capable of conveying a vast amount of information and dynamics sufficient for comprehending interoception in the top-down and bottom-up modes of causality. This model is inspired from Michael Lockwood's inversion of Hermann Minkowski's 4d light cone diagram of spacetime into an event space of the mind (see Jammer, 2000). The only difference is this model includes the hypersurface of the present in a higher dimension as a manifold containing intrinsic adaptive models of the self, simplistically depicted as a little diamond on top of a big one.

C. Turning to philosophy, the "three modes of perception" described by Whitehead are indelibly linked to interoception, the neurovisceral axis, homeostatic regulation and insular dynamics as a basis for subjective awareness (affective, emotional, self-based). Whitehead calls his theory of *actual entities* “a cell theory of actuality” (Whitehead, 1929) and yet, something like a biological basis for ‘actual entities’ linked to subjective experience remains to be fully or formally realized. Whitehead’s experience-based ‘philosophy of organism’ precipitates the notion there should be biological counterparts to the logic and causality entailed in actual entities (AE’s) and the three modes of perception originating from the distinction between consciousness and subjectivity. As highlighted in this paper:

1. The mode of "*causal efficacy*" corresponds with the "internal milieu" of the body involving homeostasis, allostasis, interoception and the neurovisceral axis. Information associated with this axis is affective, reflecting internal states of the body and intuitions from "an impressive array of signals originating in the lumen" projecting to the responsive wall of the gut (Dockray, 2003). As Bud Craig says, the main question is "*how do i feel now*" (Craig, 2009).

2. The mode of "*presentational immediacy*" in the context of interoception corresponds with the cortical target of neurovisceral (primary) afferent fibers in the insula, ACC and S1, generating a constant representation of objective physical signals from the body.

3. The mode of "*symbolic reference*" refers to the mixed mode of subjective "*re-representation*" (of objective physical signals from the body with multimodal information integration; and more-generally, of cortical information with information from the body), reinforcing a sense of body ownership and feelings accompanying each moment that relay a unique sense of "what it is like" to be you.

D. By highlighting the components of previous sections, we recognize a biological example in terms of interoception that shares goals with IIT to render subjective experience in physical terms and localize it in neural operations that can be tested in a cortical context.

Final Comments

"Speculative boldness must be balanced by complete humility before logic, and before fact." Whitehead, 1929

A major topic of discussion in neuroscience/philosophy/psychology involves the nature of consciousness, self, identity, emotion and to what extent cognition and affect play a role. Neuroscientists like Antonio and Hanna Damasio maintain that affect is at the core of conscious experience, and feelings generated from the body provide the groundwork for subjective experience. As he says, "you cannot understand feeling without consciousness, and you cannot understand consciousness without feelings" (Damasio, 2022). South African psychiatrist Mark Solms agrees: "raw feelings are the fundamental form of consciousness" (Solms, 2021). Such feelings are represented in the brain, yet indelibly attached to feelings arising from the body's internal milieu. As the Damasio's explain, "feelings are a cooperation of the brain and the body...homeostatic feelings, such as hunger, thirst, pain, or well-being, are naturally and spontaneously conscious, or they would not be of any value to us, and the knowing is really what comes out of feeling. This spontaneous knowledge helps us manage life" (Damasio & Damasio, 2022).

Additional layers of support come from the philosophy of Whitehead, who also distinguishes the role of feelings as primary to consciousness. Whitehead's 'three modes of perception' provide just the right kind of details when read into suitably generalized dynamics of interoception. What is perhaps just as remarkable are indicators of the counter-flowing top-down predictive dynamics also found in Whitehead's work (a la *propositions*, *autopoiesis*, *objective immortality, superject*). Between the two modes of causality, it is clear Whitehead focuses on sensory dynamics, yet also leaves room for the anticipatory. Reading his three modes of perception against the backdrop of interoceptive dynamics, we encounter a vivid concomitance of descriptions directly linked so well they can be identified as formal philosophical foundations for interoception.

Aristotle once remarked that "learning is not child's play; we cannot learn without pain." I disagree somewhat. Starting from a conceptual model is no more painful than a 'choose your own adventure', plus it makes all learning feel like relearning (Plato)—which *is* fun. Really, it is rare to find such a convergent topic and learning module so readily available for comprehensive packaging, as with interoception. From a learning perspective, approaching a particular field of knowledge from the vantage of a heuristic guide provides multiple jumping-off points for expanding the narrative with additional features, and for generalizing from advanced topics into more-standard curricular topics; thus, inverting the conventional approach of the learning model at university in the interdisciplinary realm of philosophy, psychology, biology, mathematics and neuroscience.

References

Allman J, Hakeem A and Watson K. Book review: Two phylogenetic specializations in the human brain. The Neuroscientist 2002;8(4):335-346.

Allman J, Watson K, Tetreault N and Hakeem, A. Intuition and autism: a possible role for Von Economo neurons. Trends in cognitive sciences 2005;9(8):367-373.

Allman J, Tetreault N, Hakeem, A, Manaye K, Semendeferi K, Erwin J, Park S, Goubert V and Hof P. The von Economo neurons in the frontoinsular and anterior cingulate cortex. Annals of the New York Academy of Sciences 2011;1225(1):59-71.

Arnhart L. The evolution of Darwinian liberalism. Journal of Bioeconomics 2015;17:3-15.

Ashby W. Principles of the self-organizing dynamic system. The Journal of general psychology 1947;37(2):125-128.

Bechara A, Damasio H and Damasio A. Emotion, decision making and the orbitofrontal cortex. Cerebral cortex 2000;10(3):295-307.

Barlassina L and Newen A. The role of bodily perception in emotion: In defense of an impure somatic theory. Philosophy and Phenomenological Research 2014;89(3):637-678.

Barrett L and Simmons W. Interoceptive predictions in the brain. Nature reviews neuroscience 2015;16(7):419-429.

Bayes T. LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S. Philosophical transactions of the Royal Society of London 1763;(53):370-418.

Bear M, Connors B and Paradiso, M, eds. Neuroscience: exploring the brain, enhanced edition: exploring the brain. Jones & Bartlett Learning; 2020.

Bechara A. The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. Brain and cognition 2004;55(1):30-40.

Bernard C. Leçons de physiologie expérimentale appliquée à la médecine: Cours de semestre d'hiver, Bailliere; 1854-1855;1.

Beul S and Hilgetag C. Towards a “canonical” agranular cortical microcircuit. Frontiers in neuroanatomy 2015;8:165.

Billman G. Homeostasis: the underappreciated and far too often ignored central organizing principle of physiology. Frontiers in physiology 2020;(11):200. https://doi.org/10.3389/fphys.2020.00200

Bossaerts P. Risk and risk prediction error signals in anterior insula. Brain Structure and Function 2010; 214:645-653.

Buchman T. The community of the self. Nature 2002;420(6912):246-251.

Buldeo N. Interoception: A measure of embodiment or attention?. International Body Psychotherapy Journal 2015;14(1).

Butti C. and Hof P. The insular cortex: a comparative perspective. Brain Structure and Function 2010;214:477-493.

Cameron O and Minoshima S. Regional brain activation due to pharmacologically induced adrenergic interoceptive stimulation in humans. Psychosomatic medicine 2002;64(6):851-861.
Cannon W. The wisdom of the body.(2nd ed.). Norton & Co; 1939.
Carhart-Harris R and Friston K. The default-mode, ego-functions and free-energy: a neurobiological account of Freudian ideas. Brain 2010;133(4):1265-1283.
Cauda F, D'agata F, Sacco K, Duca S, Geminiani G and Vercelli A. Functional connectivity of the insula in the resting brain. Neuroimage 2011;55(1):8-23.
Chanes L and Barrett L. Redefining the role of limbic areas in cortical processing. Trends in cognitive sciences 2016;20(2):96-106.
Chang C, Liu Z, Chen M, Liu X and Duyn J. EEG correlates of time-varying BOLD functional connectivity. Neuroimage 2013;72:227-236.
Chernigovskiy V and Lindsley D. Interoceptors. Trans. American Psychological Association; Ed. G. Onischenko; 1967. https://doi.org/10.1037/10531-000
Clark A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and brain sciences 2013;36(3):181-204.
Cobb J and Griffin D. Process theology: An introductory exposition. Westminster. John Knox Press; 1976.
Corbetta M, Patel G and Shulman G. The reorienting system of the human brain: from environment to theory of mind. Neuron 2008;58(3):306-324.
Craig A. How do you feel? Interoception: the sense of the physiological condition of the body. Nature reviews neuroscience 2002;3(8):655-666.
Craig A. Interoception: the sense of the physiological condition of the body. Current opinion in neurobiology 2003;13(4):500-505.
Craig A. Interoception and emotion: a neuroanatomical perspective. Handbook of Emotions 2008;3(602):272-88.
Craig A. How do you feel—now? The anterior insula and human awareness. Nature reviews neuroscience 2009;10(1):59-70.
Critchley H, Mathias C, Josephs O, O'Doherty J, Zanini S, Dewar B, Cipolotti L, Shallice T and Dolan R. Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. Brain 2003;126(10):2139-2152.

Critchley H. The human cortex responds to an interoceptive challenge. Proceedings of the National Academy of Sciences 2004;101(17):6333-6334.
Damasio A, Tranel D and Damasio H. Somatic markers and the guidance of behavior: Theory and preliminary testing. In H. S. Levin, H. M. Eisenberg, & A. L. Benton (Eds.), Frontal lobe function and dysfunction. Oxford University Press; 1991:217–229.
Damasio A, Anderson S and Tranel D. The frontal lobes. Clinical neuropsychology 1993;3:409-460.
Damasio A. Descartes' error and the future of human life. Scientific American 1994;271(4):144-144.
Damasio A. The somatic marker hypothesis and the possible functions of the prefrontal cortex. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 1996;351(1346):1413-1420.
Damasio A. Looking for Spinoza: Joy, sorrow, and the feeling brain. Houghton Mifflin Harcourt; 2003.
Damasio A. Neuroscience and the emergence of neuroeconomics. In Neuroeconomics 2009:207-213.
Damasio A. Feeling & knowing: Making minds conscious. Cognitive neuroscience 2021;12(2):65–66. https://doi.org/10.1080/17588928.2020.1846027

Damasio A and Damasio H. Homeostatic feelings and the biology of consciousness. Brain 2022;145(7):2231-2235.
Dayan P, Hinton G, Neal R and Zemel R. The helmholtz machine. Neural computation 1995;7(5):889-904.
Dayan P and Yu A. Phasic norepinephrine: a neural interrupt signal for unexpected events. Network: Computation in Neural Systems 2006;17(4):335-350.
Deen B, Pitskel N and Pelphrey K. Three systems of insular functional connectivity identified with cluster analysis. Cerebral cortex 2011; 21(7):1498-1506.

Dehaene S, Changeux J, Naccache L, Sackur J and Sergent C. Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends in cognitive sciences 2006;10(5):204-211.

Dehaene S and Changeux J. Experimental and theoretical approaches to conscious processing. Neuron 2011;70(2):200-227.

Dennett D. "Filling in" versus finding out: A ubiquitous confusion in cognitive science. In H. L. Pick, Jr., P. W. van den Broek, & D. C. Knill (Eds.), Cognition: Conceptual and methodological issues 1992:33–49. https://doi.org/10.1037/10564-002

Depraz N and Varela F. At the source of time: Valence and the constitutional dynamics of affect. Journal of Consciousness Studies 2005;2(8-10):61-81.

Dockray G. Luminal sensing in the gut: an overview. J Physiol Pharmacol 2003;54(Supplement-4):9-17.

Dunn B, Galton H, Morgan R, Evans D, Oliver C, Meyer M, Cusack R, Lawrence A and Dalgleish T. Listening to your heart: How interoception shapes emotion experience and intuitive decision making. Psychological science 2010;21(12):1835-1844.

Dunn B, Dalgleish T and Lawrence A. The somatic marker hypothesis: A critical evaluation. Neuroscience & Biobehavioral Reviews 2006;30(2):239-271.

Edelman G. Wider than the sky: The phenomenal gift of consciousness. Yale University Press; 2004.

Edlund J, Chaumont N, Hintze A, Koch C, Tononi G and Adami C. Integrated information increases with fitness in the evolution of animats. PLoS computational biology 2011;7(10): p.e.1002236.

Emerson R. "The Poet." Ralph Waldo Emerson and Margaret Fuller Selected Works: Essays, Poems, and Dispatches with Introduction; 2000:186-206.

Ernst J, Böker H, Hättenschwiler J, Schüpbach D, Northoff G, Seifritz E and Grimm S. The association of interoceptive awareness and alexithymia with neurotransmitter concentrations in insula and anterior cingulate. Social cognitive and affective neuroscience 2014;9(6):857-863.

Evrard H, Forro T and Logothetis N. Von Economo neurons in the anterior insula of the macaque monkey. Neuron 2012;74(3):482-489.

Fan Y, Duncan N, De Greck M and Northoff G. Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. Neuroscience & Biobehavioral Reviews 2011;35(3):903-911.

Friston K. Beyond phrenology: what can neuroimaging tell us about distributed circuitry?. Annual review of neuroscience 2002;25(1):221-250.

Friston K, Kilner J and Harrison L. A free energy principle for the brain. Journal of physiology-Paris 2006;100(1-3):70-87.

Friston K, Mattout J, Trujillo-Barreto N, Ashburner J and Penny W. Variational free energy and the Laplace approximation. Neuroimage 2007;34(1):220-234.

Friston K and Kiebel S. Predictive coding under the free-energy principle. Philosophical transactions of the Royal Society B: Biological sciences 2009;364(1521):1211-1221.

Friston K. The free-energy principle: a unified brain theory?. Nature reviews neuroscience 2010;11(2):127-138.

Friston K. Prediction, perception and agency. International Journal of Psychophysiology 2012; 83(2):248-252.

Friston K, Schwartenbeck P, FitzGerald T, Moutoussis M, Behrens T and Dolan R. The anatomy of choice: active inference and agency. Frontiers in human neuroscience 2013;7:598.

Friston K, Levin M, Sengupta B and Pezzulo G. Knowing one's place: a free-energy approach to pattern regulation. Journal of the Royal Society Interface 2015;12(105):20141383.

Frost R. The Hartford Courant, Frost, Poet, Is Honored On Birthday: New Englander, 80, Views World as Too Crowded and Hurried; Hartford, Connecticut. March 26 1954; Column 1:14.

Garfinkel S and Critchley H. Interoception, emotion and brain: new insights link internal physiology to social behaviour. Commentary on: “Anterior insular cortex

mediates bodily sensibility and social anxiety" by Terasawa et al. 2012. Social cognitive and affective neuroscience 2013;8(3):231-234.
Goldstein D and Kopin I. Evolution of concepts of stress. Stress 2007;10(2):109-120.
Gregory R. Perceptions as hypotheses. Philosophical Transactions of the Royal Society of London. B, Biological Sciences 1980;290(1038):181-197.
Grenham S, Clarke G, Cryan J and Dinan T. Brain–gut–microbe communication in health and disease. Frontiers in physiology 2011;2:94.
Gross J. The emerging field of emotion regulation: An integrative review. Review of general psychology 1998;2(3):271-299.
Gu X and FitzGerald T. Interoceptive inference: homeostasis and decision-making. Trends Cogn Sci 2014;18(6):269-70.
Gu X, Hof P, Friston K and Fan J. Anterior insular cortex and emotional awareness. Journal of Comparative Neurology 2013;521(15):3371-3388.
Hale M, Shekhar A and Lowry C. Stress-related serotonergic systems: implications for symptomatology of anxiety and affective disorders. Cellular and molecular neurobiology 2012;32:695-708.
Haken H and Haken H. An introduction: nonequilibrium phase transitions and self-organization in physics, chemistry and biology. Synergetics: Introduction and Advanced Topics 2004:1-387.
Harrison N, Gray M, Gianaros P and Critchley H. The embodiment of emotional feelings in the brain. Journal of Neuroscience 2010;30(38):12878-12884.
Helmholtz H. Concerning the perceptions in general, 1867. In: W. Dennis (Ed.), Readings in the history of psychology. Appleton-Century-Crofts; 1948:214–230. https://doi.org/10.1037/11304-027
Hohwy J. The predictive mind. OUP Oxford; 2013. https://doi.org/10.1093/acprof:oso/9780199682737.001.0001
Holzer P. Interoception and gut feelings: Unconscious body signals' impact on brain function, behavior and belief processes. Processes of believing: The acquisition, maintenance, and change in creditions 2017:435-442.
James W. The principles of psychology, Vol. 1. Henry Holt and Co. New York; 1890. https://doi.org/10.1037/10538-000
Jammer M. Einstein and religion: Physics and theology. Princeton University Press; 2000.
Kauffman S. The origins of order: Self-organization and selection in evolution. Oxford University Press, USA; 1993.
Kirk U, Downar J and Montague P. Interoception drives increased rational decision-making in meditators playing the ultimatum game. Frontiers in neuroscience 2011;5:49.
Knill D and Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. TRENDS in Neurosciences 2004;27(12):712-719.
Krauss P and Maier A. Will we ever have conscious machines?. Frontiers in computational neuroscience 2020;14:116.
Kuo J and Linehan M. Disentangling emotion processes in borderline personality disorder: physiological and self-reported assessment of biological vulnerability, baseline intensity, and reactivity to emotionally evocative stimuli. Journal of abnormal psychology 2009;118(3):531.
Lange C. Om sindsbevaegelser; et psyko-fysiologisk studie; Lund.; 1885.
Lango J. Whitehead's Ontology. State University of New York Press; 1972.
Lazarus R. Progress on a cognitive-motivational-relational theory of emotion. American psychologist 1991;46(8):819.
Lee S. A Copernican approach to brain advancement: The paradigm of allostatic orchestration. Frontiers in Human Neuroscience 2019;13:129.
Leikind M. Source Book of Medical History. Compiled with Notes by Logan Clendening, Professor of the History of Medicine, University of Kansas. (New York: Paul B. Hoeber, Medical Book Department of Harper and Brothers 1942:xiv;685).

Lutz A and Thompson E. Neurophenomenology: integrating subjective experience and brain dynamics in the neuroscience of consciousness. Journal of consciousness studies, 2003;10(9-10):31-52.

Maclachlan D. Whitehead's Theory of Perception. Process studies 1992;227-230.
Marstaller L. Towards a Whiteheadian neurophenomenology. Concrescence-The Australasian Journal of Process Thought 2009:10.
Mayer E. Gut feelings: the emerging biology of gut–brain communication. Nature Reviews Neuroscience 2011;12(8):453-466.
Mesulam M and Mufson E. Insula of the old world monkey. III: Efferent cortical output and comments on function. Journal of Comparative Neurology 1982;212(1):38-52.
Montague P Lohrenz T. To detect and correct: norm violations and their enforcement. Neuron 2007;56(1):14-18.
Mesulam M and Mufson E. Insula of the old world monkey. Architectonics in the insulo-orbito-temporal component of the paralimbic brain. Journal of Comparative Neurology 1982;212(1):1-22.
Nicolis G and Prigogine I. Self-organization in Nonequilibrium Systems. From Dissipative Structures to Order through Fluctuations 1977:91.
Ohira H. Interoception and decision-making. Nihon Shinkei Seishin Yakurigaku Zasshi= Japanese Journal of Psychopharmacology 2015;35(1):11-18.
Oizumi M, Albantakis L and Tononi G. From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. PLoS computational biology 2014;10(5):p.e1003588.
Ondobaka S, Kilner J and Friston K. The role of interoceptive inference in theory of mind. Brain and cognition 2017;112:64-68.
Parvizi J and Damasio A. Consciousness and the brainstem. Cognition 2001;79(1-2):135-160.
Peirce C. Five Hundred and Eighty-First Meeting. Monthly Meeting; On the Natural Classification of Arguments. In Proceedings of the American Academy of Arts and Sciences 1867; (7):261-287.
Picard F and Friston K. Predictions, perception, and a sense of self. Neurology 2014;83(12):1112-1118.
Popper K. Science as falsification. Conjectures and refutations 1963:33-39.
Preuschoff K, Quartz S and Bossaerts P. Human insula activation reflects risk prediction errors as well as risk. Journal of Neuroscience 2008;28(11):2745-2752.

Prinz J. Gut reactions: A perceptual theory of emotion. Oxford University Press; 2004.
Rhee S, Pothoulakis C and Mayer E. Principles and clinical implications of the brain–gut–enteric microbiota axis. Nature reviews Gastroenterology & hepatology 2009;6(5):306-314.
Rilling J, King-Casas B and Sanfey A. The neurobiology of social decision-making. Current opinion in neurobiology 2008;18(2):159-165.
Roitblat H. Representations and processes in working memory. In: Animal Cognition; Psychology Press; 2013:189-206.
Rosas F, Mediano P, Biehl M, Chandaria S and Polani D. September. Causal blankets: Theory and algorithmic framework. In International workshop on active inference; Cham: Springer International Publishing; 2020:187-198).
Saper C. The central autonomic nervous system: conscious visceral perception and autonomic pattern generation. Annual review of neuroscience 2002;25(1):433-469.
Schachter S and Singer J. Cognitive, social, and physiological determinants of emotional state. Psychological review 1962;69(5):379.
Seth A. Interoceptive inference, emotion, and the embodied self. Trends in cognitive sciences, 2013;17(11):565-573.
Seth A. The cybernetic Bayesian brain: from interoceptive inference to sensorimotor contingencies. In Open MIND (eds Windt JM, Metzinger T); Frankfurt, Germany: MIND Group; 2015:9–24.
Sherrington C. The Integrative Action of the Nervous System, by Sir Charles Sherrington... With a New Foreword by the Author and a Bibliography of His Writings. The University Press; 1948.

Solms M. The hidden spring: A journey to the source of consciousness. Profile Books; United Kingdom; 2021.
Stavros B. Empedocles and Neurosciences. Encephalos 2014;51:66-78.

Sterling P and Eyer J. Allostasis: A new paradigm to explain arousal pathology. In S. Fisher & J. Reason (Eds.), Handbook of life stress, cognition and health; John Wiley & Sons; 1988;629–649.

Sterzer P and Kleinschmidt A. Anterior insula activations in perceptual paradigms: often observed but barely understood. Brain Structure and Function 2010;214(5-6):611-622.

Tajadura-Jiménez A and Tsakiris M. Balancing the “inner” and the “outer” self: Interoceptive sensitivity modulates self–other boundaries. Journal of Experimental Psychology: General 2014;143(2):736.

Tsakiris M, Jiménez A and Costantini M. Just a heartbeat away from one's body: interoceptive sensitivity predicts malleability of body-representations. Proceedings of the Royal Society B: Biological Sciences 2011;278(1717):2470-2476.

Tononi G and Koch C. The neural correlates of consciousness: an update. Annals of the New York Academy of Sciences 2008;1124(1):239-261.

Tononi G. What is the Essence of Consciousness? [Video]. Closer to Truth. YouTube. March 30, 2021. Accessed 15 May 2023.

Touroutoglou A, Hollenbeck M, Dickerson B and Barrett L. Dissociable large-scale networks anchored in the right anterior insula subserve affective experience and attention. Neuroimage, 2012;60(4):1947-1958.

Turrigiano G. Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. Trends in neurosciences 1999;22(5):221-227.

Vaitl D. Interoception. Biological psychology 1996;42(1-2):1-27.

Van Den Heuvel M and Sporns O. Rich-club organization of the human connectome. Journal of Neuroscience 2011;31(44):15775-15786.

Verdejo-Garcia A, Clark L and Dunn B. The role of interoception in addiction: a critical review. Neuroscience & Biobehavioral Reviews 2012;36(8):1857-1869.

Watson K, Jones T and Allman J. Dendritic architecture of the von Economo neurons. Neuroscience 2006;141(3):1107-1112.

Watson K, Matthews B and Allman J. Brain activation during sight gags and language-dependent humor. Cerebral cortex 2007;17(2):314-324.

Weber M and Weekes A. eds. Process approaches to consciousness in psychology, neuroscience, and philosophy of mind. State University of New York Press, Albany; 2009.

Whitehead A. Symbolism, its meaning and effect. Macmillan, United Kingdom; 1927.

Whitehead A. Process and reality: An essay in cosmology. Cambridge University Press; 1929.

Wiese W. Perceptual presence in the Kuhnian-Popperian Bayesian brain. In: Open MIND (eds Windt JM, Metzinger T). Frankfurt am Main: MIND Group; 2014.