**1. Introduction to Emotions: An ideological, historical, and transactional overview**

Mirela Ramacciotti, Federal University of São Paulo (UNIFESP)

Av. Prof. Mello Moraes, 1721, São Paulo, Butantã, Cidade Universitária

CEP 05508-030 Brazil

**Abstract**

Why do humans need to move outwards? Needless to say, it is a rhetorical, semantic question but one that lies at the heart of emotions. As the world has grown in depth and complexity, our participation in it, along with the progressive inclination to engage with others sharing the same spatial context over time, has grown markedly. In the field of emotions, it meant a work that is vast, complex, and still in development. Thus, our work in this chapter is restricted to providing **reference points** in a review meant to account for how the study of emotions developed (i) ideologically, i.e., what notions steered taxonomies; (ii) historically, i.e., how main contributions have shaped our understanding of emotions across time; (iii) transactionally, i.e., how contributions from multiple researchers and fields have shaped current understandings of emotions. The first section looks at reference points which are crucial when we need definitions that allow for comparing and contrasting concepts, like those of affect, feelings, and emotion. In the second section, our focus will be on neuroscientific perspectives that have reshaped how we view emotions. In the third section we recall some key figures to delve into how we currently understand emotions based on their insights. We expand on their contributions to examine what thresholds and breakthroughs they have brought to the field of emotions. In a field that has ebbed and flowed with contributions along almost two centuries of intensive scientific pursuit, it is befitting that we emulate the same movement to convey the intricacies that a thorough understanding implies. Let us start with how emotions became the construct they are today, i.e., their epistemic origins.

Keywords: Historical background; Theoretical emotion models; Neuroscience

**1.1 Introduction**

In the Western world, around the 16th century, the word “emotions” entered our vocabulary. Based on the Latin word *emovere*, from *e-* (variant of *ex-*) meaning ‘out’ joined with *movere* meaning ‘move’, the term was used to refer to a public disturbance, as literature in the French language attests (Dixon, 2012). However, in the Eastern world, in China more specifically, origins date back to the early sixth to fifth centuries BCE (Seok, 2021). Interestingly, Chinese origins of emotions are linked to that of *qing* (情), an ideogram that semantically encompasses – among other uses and transformations over time – ontological meanings of essence, nature, reality *en par* with psychological meanings of disposition, desire, and feelings that emerge **genuinely** (Seok, 2021). From its inceptions, emotions in Chinese philosophy are fundamentally based on interactive dispositions, thus involving genuine perceptions, motivations and affections of the self in relation to others (Seok, 2021).

Back in the western world, the spelling in French – *émotion* – derived from the word *émouvoir* meaning ‘excite’. The French philosopher Descartes tried to incorporate the term into his works as a substitute for the term passion (Dixon, 2012). However, it took more than a century, and other contributors – David Hume and Adam Smith – for emotions to figure as a general term denoting mental agitation (Dixon, 2003; 2012). In the 17th century, the term entered the English language through a translation of Michel de Montaigne’s essays (Dixon, 2012).

In both (English and French) linguistic traditions, the term 'emotion' historically denoted a form of physical perturbation—manifesting as public commotion, bodily agitation of organic matter, and the animation of entities, whether human or non-human (Dixon, 2012). During the 18th century, emotion denoted a combination of bodily agitation and mental feelings (Dixon, 2012). It was in the 19th century, though, that emotions definitely entered our language as referring to appetites, desires (or passions) and affections (or sentiments) in humans. This feat was achieved by the moral philosopher and professor, Thomas Brown, during his lectures on the science of the mind (Dixon, 2003; 2012). Since then, emotions have designated all-encompassing feelings (and their states), sentiments, pleasures, pains and affections in a term that went from folk usage to the scientific realm (Dixon, 2012). This turned the work of those involved in understanding and researching emotions into a herculean task. Arguably, hitherto have come the multitude of theories and conceptions that has populated the field ever since.

Emotion as a keyword became known through the works of William James due to his usage of the term in the article “What Is an Emotion?” (Dixon, 2012; James, 1884). James brought to the fore the idea that emotions as feelings are not static states but rather dynamic, i.e., subject to change. This happens according to changes in perceptions of physiological conditions, meaning autonomic and motor functions (Scarantino & de Souza, 2021). Perhaps the following phrase best summarizes this key notion: “our feeling of [bodily] changes as they occur is the emotion*”* (James, 1884, pp 189–190).

Hence from the Western tradition, it carries a lot of weight, semantically speaking, in thinking that emotions are what makes us connect to the outside world – the moving outwardly of our initial quest – generating some disturbance or change thereby denoting certain excitement. In that sense, it is not dissociated with the Chinese philosophical conceptualization of emotions for they are taken as holistic affects that include senses, thinking, and feelings in constant interaction with the outside world (Seok, 2021). In short, emotions constitute communication heuristics[[1]](#footnote-1); they provide shortcuts for internally generated messages that are channelled out via facial configurations. If there is universality in them, it should be taken as genuine expressions of individualized, constantly reconstructed messages which depend on others to be construed at each moment in time.

Thus, emotion expressions communicate instances, states, and experiences, much like language does (Kalateh et al., 2024). Henceforth, emotions should not be taken as universally pre-set, deterministic phenomena, as many still believe, and most perpetrate, based on identification of standardized facial expressions. In addition, the notion that emotions communicate bodily states without the need for words does not mean that words should not be considered when emotions are created and/or communicated (Barrett, 2017a; 2020). Emotions, by virtue of their communicative function, serve to convey internal states both to oneself and to others. In doing so, they integrate elements of creation—understood as the recombination of fundamental components—and of communication, which presupposes the transmission of meaning between a sender and a receiver grounded in shared conceptual frameworks (Berlo, 1960 in the communication model known as source-message-channel-receiver, or SMCR).

Therefore, language – meaning the use of words – becomes a crucial element when the concept of emotion gets formed in one’s mind (Barrett, 2017a; 2020). Further, language is also necessary to enable recognition by another mind that sits at the receiver end and has to construe what a certain emotion entails in the social reality shared by the sender and receiver (see Figure 1).

**Language**

**Figure 1**: The SMCR model (Berlo, 1960) applied to an overarching understanding of emotions based on meaning-making dependent on language. Source: Author.

Of note, emotions imply mutual understanding and carry with them a repertoire of words that generate widely shared concepts. Such concepts allow for categorizations underlying constructs that become shared. These constructs enable interpretations which tend to be shared, reinforced and reiterated. When considering universality that emotions entail, appreciation should be turned not to their end means, i.e., their expression, but rather to their origins, i.e., how emotions got to be what we understand them to be. This process happened through constructs and concepts acquired over time as individuals developed and interacted with the contexts in which they found themselves. As we shall later appreciate in greater detail, recent research on brain encoding of emotion concepts showed how representations of emotion concepts found stability by late childhood and synchronization between individuals in adolescence (Camacho et al., 2023). This stability is intrinsically related to statistical learning. As perceptions shift over time, the way emotions develop and are learned since infancy happens through extraction of environmental patterns that are subject to change (LoBue and Ogren, 2022). These patterns sustain the regularities of emotion display in different social contexts (Wang, Lu,, & Wu, 2025). Critically, this extraction is regarded as an implicit learning mechanism, meaning that consciousness or awareness may not be implicated in the process (Conway, 2020).

This upends the way that we discuss emotions. Emotions do rely on biological mechanisms engendered in our species, but they are shaped by socioemotional feedback, which is enhanced by words, that change and differ (from language to language). Words enable us to form the mental concepts that are externalized – sometimes via facial configurations - in cultural contexts. Such are the shared spaces where emotions acquire a transactional nature due to the co-construction that shapes our realities from time to time (Barrett, 2017a; 2017b; 2020). Enlarging this context in geographical terms, as in westernized-only versus other-worldly reports of shared constructs, may shift how emotions are configured. Cultural differentiation also matters in how physiological aspects are viewed and narrated. For instance, in studying an African tribe’s account of emotionally laden experiences, participants were found to rely much more on concrete and contextual factors than their westernized counterparts, who relied more heavily on thoughts and feelings (Hoemann et al., 2023). There is more to emotions than subjective feelings and other internal mental states. In summary, biology and language are universal pathways that enable emotions which are construed and modulated by cultural and situated constraints.

Thus, taking emotions for expressions that are universal and have an evolutionary significance to justify automation for emotion detection might conflate ideas that deserve a more careful treatment. For instance, pushing forward the notion that emotions are universal to justify shortcuts in emotion recognition means overlooking important differences that cultures hold in making facts out of the meaning that emotions carry (Mesquita, Boiger, & De Leersnyder, 2017; Mesquita & Walker, 2003). Further, culture shapes perceptions to the point of running counter to the evolutionary perspective of emotions as universal features impervious to cultural modulation (Crivelli et al., 2016; 2017; Gendron et al., 2018; 2020) or to the situation or goal one may hold from time to time (Hoenemann et al., 2023).

Treating emotions in a careful way implies digging into their basic foundations, those that takes us from the simple - yet prescient - semantically rhetorical question asked in the beginning – to the cardinal points in the study of emotions as definitions that are very hard to come by (Adolphs, Mlodinow, & Barrett, 2019; de Waal & Andrews, 2022; Dixon, 2012; Izard, 2010a; 2010b; LeDoux 2021). Treating emotions this way may prevent us from misconstruing concepts, such as smiles equating joy and tears equating sadness.

Emotion, as an area of knowledge, undergoes constant progress to account, for example, for physiological activity (Hoemann et al., 2023) and multiplicity in facial movements (Dúran & Fernández-Dols, 2021) that point towards the importance of embodiment, culture, and language (the cardinal points mentioned in the previous paragraph). Both topics – physiology and facial movements - signal endeavours to uncover bodily and mental events that exemplify and reflect neural representations in more ways than one (Westlin et al., 2023). They also encompass physical and bodily properties that enable cellular ensembles - in a many to one fashion (Westlin et al., 2023) - of neural patterns. For these collections to become meaningful prescriptions for actions – or emotions - we need a brain that detects and interprets interoceptive signals; a mind that develops them into concepts; and other people that co-construct a social reality where emotions become “sources of wisdom” when displayed (Barrett, 2017a, p. 246). Doing so may allow us to comprehend what emotions entail.

And while we have opted for the use of “humans” focusing attention on our species, it does not follow that other animals do not have emotions. It is just a necessary resource to limit this overview to the inherent complexity – originating from multiple causality and implying variable instead of normative treatments - that contemplating emotions in our species entails. It also entails treating emotions as conscious precepts that allow for differentiated assessment (Malezieux, Klein, & Gogolla, 2023).

**1.2. Emotions: Notions that steered taxonomies**

In the first half of the 20th century, Donald O. Hebb (1946) ascertained that emotions could be referenced as “summary descriptions and predictions of behavior” (op. cit, p. 89). He also advanced that emotions are “inferred special states (...) of changed responsiveness” (Hebb, 1946, p. 104). Presciently, Hebb touched on some aspects that have proved relevant to how the study of emotions has developed, especially in the present century. Let us examine them in turn.

A summary, as stated by Hebb, presupposes a combination of primary elements, those that provide the main facts or ideas about something (as stated by the Cambridge Dictionary, online version[[2]](#footnote-2)). To pave our understanding, emotions may be constituted of distinct affects, the primary elements, that combine to give a main idea about an emotional state. This indeed seems to be the current direction in the field (Hoemann et al., 2023). An inference, as stated by Hebb, presupposes that a given result comes from a most probable (inferred) cause. Hence, emotional states may be regarded by each individual subjectively based on the most likely cause(s) available (Barrett et al., 2019). According to Hebb, a special physiological state arises in organisms striving for constancy or homeostasis within dynamic environments. This state presupposes that the brain continuously regulates metabolic states by integrating signals that indicate changes requiring intervention (Barrett, 2017a; 2020). Thus, emotions may be taken as concepts formed over time that attribute meaning to interoceptive, sensory and musculoskeletal signals (Barrett, 2017a; 2020). According to this viewpoint, these concepts emerge as an integrative experience that spans over sensory modalities, i.e., it is supramodal (Barrett, 2017b). Another viewpoint attributes such meaning as independent of the sensation evoked, i.e., an amodal experience (Kryklywy et al., 2020). In section 1.4 we further explore what these different viewpoints mean to emotions. Returning to Hebb's influential work, he discussed the prediction of behaviours as contingent upon contextual factors. Therefore, the specific situation, cultural background, and manner of emotional expression should be incorporated into a comprehensive understanding of emotions and the discourse surrounding them (Barrett, 2017a; 2020; Hoemann et al., 2023; Immordino-Yang; Yang, & Damasio, 2017; Saxbe et al., 2013).

Defining a construct is necessary for the application of a scientific treatment. This is especially relevant for emotions, a field with a long-suffering history of a lack of consensus (Gendron, 2010; Gendron & Barrett, 2009). Over time, different strands have merged into clusters based on how characteristics were combined. Gendron and Barrett (2009) presented distinct researchers whose work can be understood in the following succinct way: (i) The *basic emotions approach* combines theories and works that have taken emotions to be innate and universal, reflexive and homologous, composing a fixed set of recognizable facial expressions largely independent of culture and language; (ii) The *appraisal approach* combines theories and works that claim emotions are intentional states that presuppose a cognitive mechanism which emerges via meaningful interpretation of situational events; and (iii) The *psychological constructionist approach* takes emotions as subjective, individual experiences constructed or inferred based on psychical compounds pertaining to different mental states which are combined and generated in varied ways dependent on cultural, contextual, and linguistic influences. Figure 2 summarizes how these approaches share an evolving interrelatedness yet preserve their distinctness focused on how emotions are characterized.

Diagrama

O conteúdo gerado por IA pode estar incorreto.

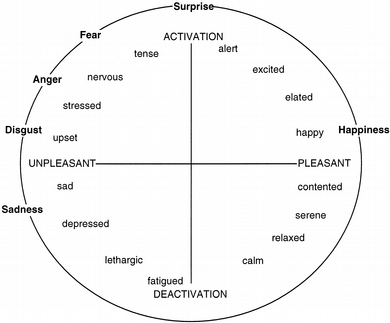
**Figure 2**: Main approaches to how emotions are characterized. Source: Author based on Gendron & Barrett (2009).

Given such a definition, emotions have been distinguished over time based on their independent qualities. This had a precursor in the works by Wilhehm Wundt, as related by Gendron and Barrett (2009). According to them, emotions are distinct from simple feelings or momentary affective states. This is a relevant distinction and can exemplify how emotions become a complex area. Whereas some authors take emotions to be synonymous with **affect** (*e.g.*. Kuppens, 2015), others distinguish them (*e.g.*, Barrett, 2017a; 2020) by the point of complexity (Reitsema et al., 2022). Concerning affect, operationally defined as the conscious feeling of experiences lived, core elements are (i) valence (pleasant and unpleasant); (ii) arousal (arousing and subduing); and (iii) intensity (straining and relaxing). But it was with appraisal, meaning how one interprets a situation, that feelings of attraction or aversion as distinctive processes underlying emotions came to the fore (Arnold, 1960). Recognizing that the same stimulus can evoke different emotions in the same person depending on timing and context, the field advanced by understanding that actions such as withdrawal or approach are driven not by events themselves, but by how individuals appraise these events and the emotions that result. Arnold (1960) devised a three-dimensional structure for appraisal depending on an internal, individualized analysis of the ensuing circumstances that could be ascertained as good or bad, present or absent, and easy to attain or avoid (Scarantino & de Souza, 2021).

Further, when emotions were taken as constructs based on affective states, patterns began to emerge. Such patterns can be different between individuals (Barrett, 2017a; 2020) for reasons like culture and language as referenced before, but also within the same individual (Fisher, Medaglia, & Jeronimus, 2018) in any single day. Understanding how emotions change means conceptualizing them in terms of dynamics that fluctuate over different time periods (as in hours, days, years, and developmental periods) over the life course (Barrett et al., 2016; Houben et al., 2015; Reitsema et al., 2022). Emotional dynamics encompass the intensity (strength), variability (dispersion), inertia (temporal persistence) and instability (temporal dependent fluctuations) that a single emotional concept may have and also the augmentation (increase) and/or blunting (decrease) together with differentiation or granularity (specificity) multiple emotion concepts may display over time (Reitsema et al., 2022).

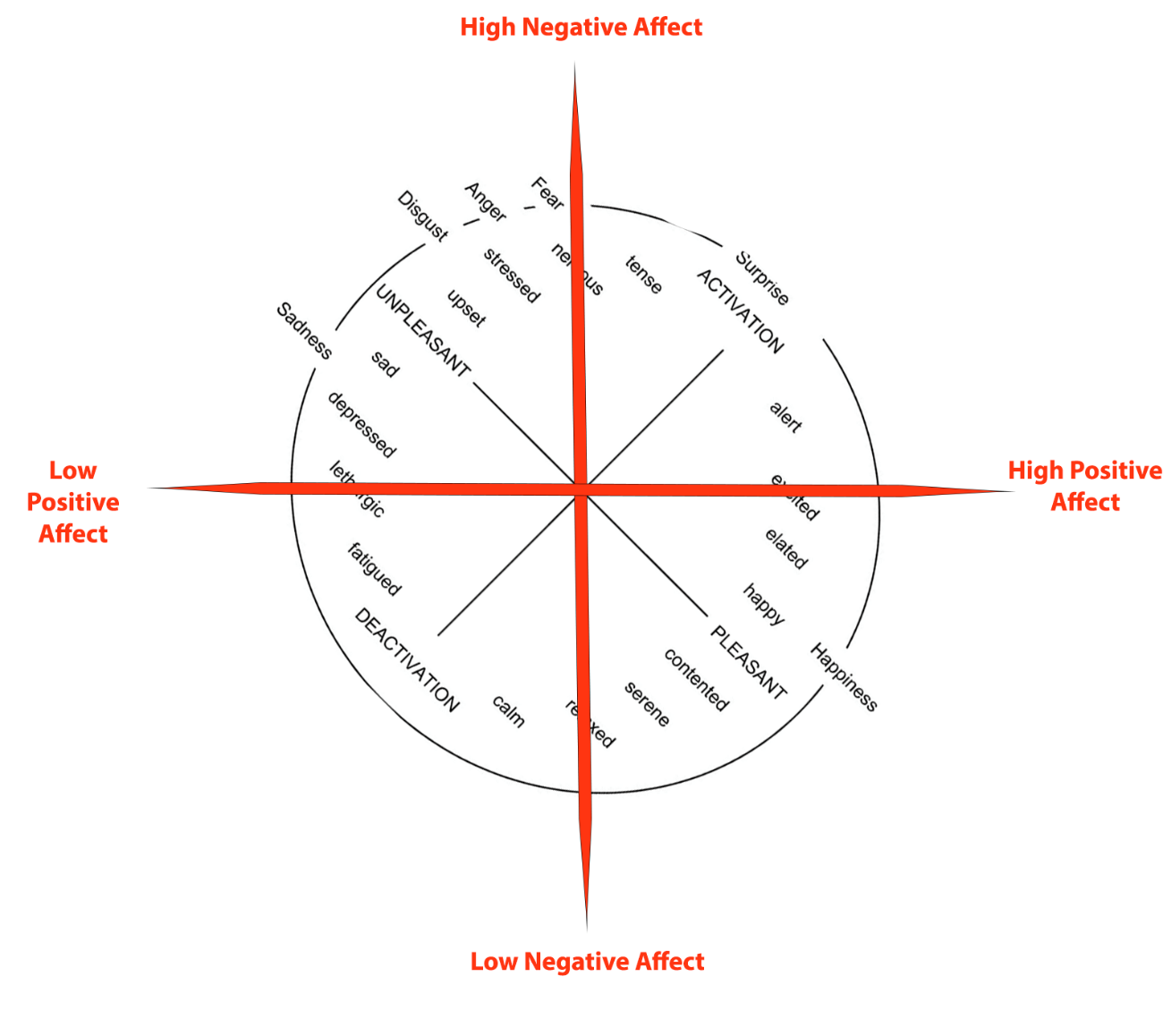
In the face of the inherent complexity of tackling a topic that has failed to reach a consensus and that has so many variables to account for, here we will proceed with an analysis of attributes considering that: (a) emotions may be positive or negative; and while some have taken them to vary along a continuum with paired opposites, such as ecstasy and grief (Plutchik, 2001), others have taken them to be discrete, distinctive – although interrelated – elements in a contained set, such as Ekman’s seven basic emotions (Ekman, 1992); and (b) emotions are the precursors of motivated behaviours, such as when one runs away because they get afraid of unusual sounds like the roar of a bear (James, 1884). Taken together, those two attributes have given rise to various theoretical taxonomies of emotions, a few of which we have singled out to be succinctly analysed next.

When emotions are considered in a continuum, dimensional circumplex models have been created to organize them. Russell (1980) laid out emotion/affect in a bidimensional activation model where arousal (high or low levels of physiological excitation on a *y axis*) and valence (positive and negative activation on a hedonic scale or *x axis*) are used to map emotions in a cartesian plane which vary in intensity according to how far they are from the centre (see Figure 3).



**Figure 3**: Russell’s (1980) Circumplex Model.

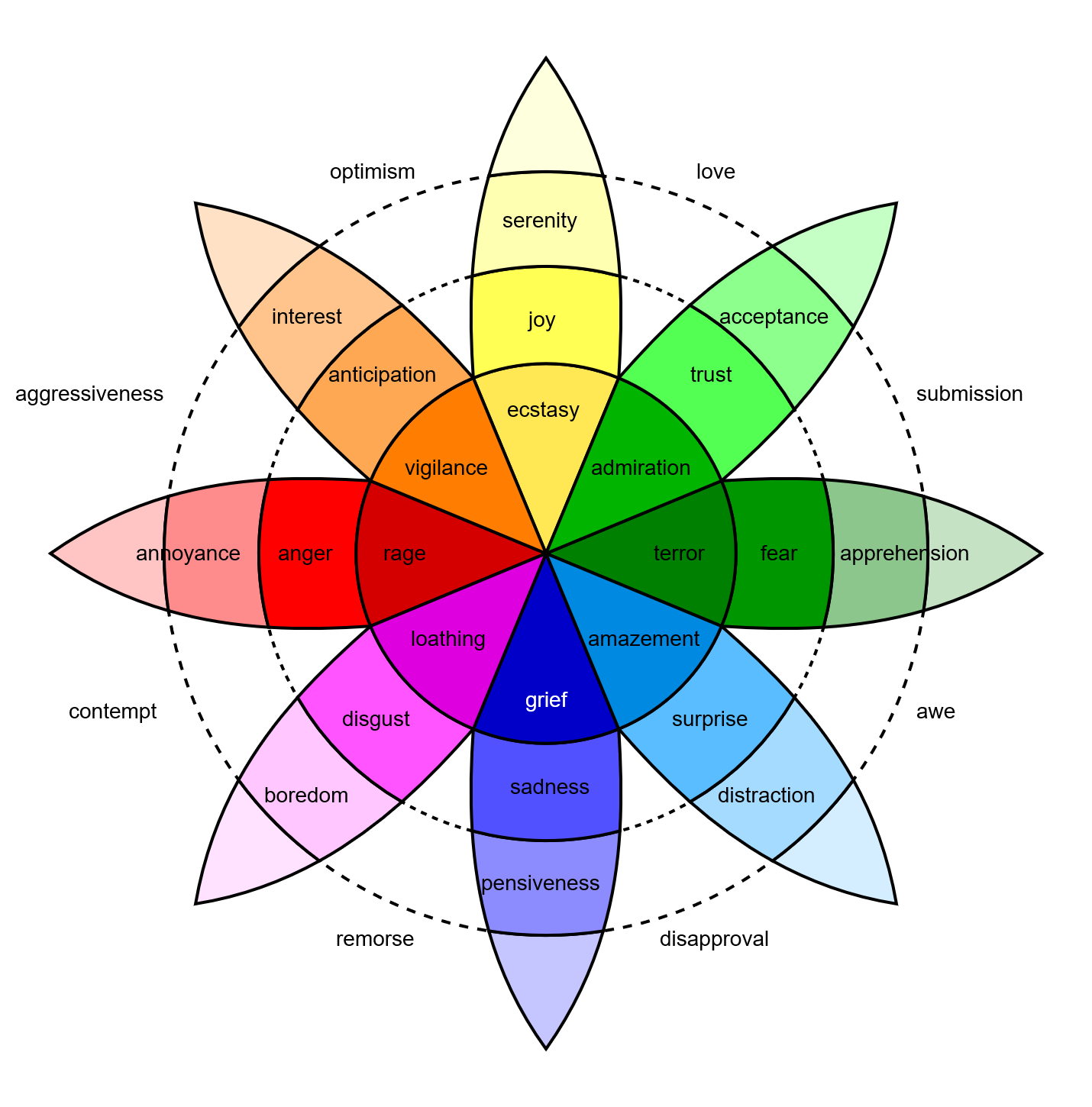
Since then, the model has been further developed by Watson and Tellegen (1985) who rotated the original circumplex model by 45 degrees (see Figure 4). That allowed them to add two new dimensions termed High Positive Affect and High Negative Affect. The result was a set of eight octants spread out over four dimensions: (1) pleasantness/unpleasantness; (2) engagement/disengagement; (3) positive affect; and (4) negative affect. Later Watson et al., (1999) termed Positive Affect and Negative Affect as Positive Activation and Negative Activation. This was done to emphasize that such terms represent highly activating emotions.



**Figure 4**: Watson and Tellegen’s (1985) 45° rotation of Russell’s (1980) circumplex model.

In 1999, Russell and Barrett defined core affect as a conscious process that can be reached via pleasant/unpleasant feelings (x axis) and activation (y axis). The novelty of their model lay with “prototypical emotional episodes” that could happen longitudinally and might be categorized (as anger, fear, shame, etc.).

In 2001, Plutchik presented a three-dimensional model based on a coloured wheel. The vertical axis of the cone shaped model accounts for intensity while the circular dimension represents similarity and builds on four opposing pairs – ecstasy-grief; admiration-loathing; terror-rage; amazement-vigilance – of the eight primary emotion dimensions that vary in degree of arousal (see Figure 5).



**Figure 5**: Pluchik’s Wheel of Emotions. Source: Created by Machine Elf 1735 - Own work, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=13285286>

Opposing a Cartesian view, Barrett proposed in 2006 that emotions are constructed concepts. Such concepts would be steered by core affects, or neurophysiological states, which build upon valence (pleasure or displeasure) representing subjectivity, and upon arousal (or the amount of uncertainty when brains try to predict the meaning of the signals they get from metabolic states). Affects would activate or deactivate concepts in instances rather than categorically. That would explain why emotions could vary according to the situation and the individual as they are dependent on the context one is in. We will come back to this notion in greater depth in Section 1.4.

When emotions are taken as discrete categories - or labels – they tend to be considered as innate and universal, therefore serving an evolutionary purpose that builds on primary elements. These elements, in turn, may be combined to give rise to more elaborate states. Paul Ekman (1992) is credited with work on a set of seven basic emotions, although stating that “no emotion exists as a single affective or psychological state” (<https://www.paulekman.com/universal-emotions/>). Together with other colleagues, Ekman theoretically and experimentally developed, along two decades of research, the notion of seven core emotions (joy, anger, fear, surprise, sadness, disgust, and contempt) taken as automatic appraisals that transcend cultural differences. Most of these discrete categories were also taken as innate, but termed “affects” by Sylvan Solomon Tomkins, a mentor to Ekman. Tomkins, in a series of four volumes (1962, 1963, 1991, 1992), defined the affect system as consisting of dyads (interest-excitement, enjoyment-joy, surprise-startle, distress-anguish, anger-rage, fear-terror, shame-humiliation), disgust [to noxious tastes] and revulsion [to noxious odours] to account for possible variations. Another disciple of Sylvan Tomkins, Carroll Izard, provided an enlarged set of ten discrete emotions (fear, anger, shame, contempt, disgust, guilt, distress, interest, surprise, and joy) which he insisted were distinct and innate (Izard, 1977). At a time when emotions were indiscriminate, Izard defended the idea that emotions served the purpose of adaptation and could be expressed distinctively via facial feedback.

Except for the notion developed by Barrett (2006), theories developed under both models – discrete or dimensional – subscribe to the idea that emotions are bound to, or emerge from, a primary and innate set, recognizable across the human species, i.e., they are not culturally responsive. This leads to the notion that they have developed alongside our evolution as mammalians. Hence, we proceed now to a brief exploration of historical beginnings with ideas put forth over time.

**1.3 Emotions: Brief overview of main contributions across time**

In recollecting some milestones for the understanding of emotions, we intend to give their main ideas spread over the past couple of centuries in tandem with the objectives of their work. For a dedicated historical account, we recommend readers to the works of Dalgleish (2004), Barrett and Lindquist (2008), Gendron and Barrett (2009) and Gendron and Barrett (2017, especially for faces).

The field of affective neuroscience began over 150 years ago (Graeff, 2018). It is taken generally as initiating with Charles Darwin’s book *The Expression of Emotions in Man and Animals* (Darwin, 1872). In this publication, initially conceived as a chapter of a previous book (*Descent of Man* published in 1871), Charles Darwin proposed a biological, evolutionary trait untethered from culture and learning (Castilho, 2021). In this proposition, basic emotions would be constituted into modules generating specific behaviours manifested as joy, fear, sorrow, and anger (Darwin, 1872). Darwin’s theoretical proposal was based on the observation of animals, in general, and on the correlation of these observations with human behaviours (Graeff, 2018). In his proposal, Darwin generated a postulate; that emotions would have an adaptive value, contributing to a natural selection of the best set of emotions, understood as reactions, which would allow humans to better adapt to environments - therefore preserving the individual and the species – in order to lead to better social communication. Darwin also postulated that such evolutionary sets of reactions would be homologous, i.e., bear correspondence between species (Graeff, 2018).

Darwin’s theory inspired several followers such as Walter Rudolf Hess, Nobel laureate in Physiology or Medicine in 1949 for having made the link between pre-organized neural networks in the central nervous system (CNS) and basic emotions, such as fear. By working with animal models, Hess stimulated the medial hypothalamus of cats, the subcortical area of the diencephalon, and observed defence responses similar to those presented to natural threats (Hess & Brügger, 1981). In the experimental paradigm, the cat manifested the same behaviour (fear) as if it were facing an aggressive dog. Through this paradigm, Hess claimed to have furnished evidence of the Darwinian theory of the expression of emotions, emphasizing the role of the hypothalamus in coordinating the functions of internal organs. Hess shared the Nobel Prize that year with neurosurgeon Antônio Egas Moniz for the surgery he had performed – a lobotomy of the prefrontal cortex – in order to deal with certain psychoses or mental disorders. Dr. Moniz was the teacher of the neuroscientist Antonio Rosa Damasio, an important researcher in our current knowledge of emotions who we will return to, in greater depth, in the next section.

In the last quarter of the 19th century, William James (1884), in the United States, and Carl G. Lange (1885), in Denmark, held coincidental – yet separately proposed - theories about how emotions evolve. To both of them, organic sensations were responsible for physiological responses - vasomotor or circulatory according to Lange (1885) and visceral according to James (1884) - and behavioural responses. Their views were so contingent on physiological manifestations that, should they be removed from the emotional experience, nothing would remain according to their propositions (Cannon, 1927).

Thus, to James and Lange, sensations – or bodily changes - came before the individual perception experienced subjectively as an emotion. Their joint views of emotions as organic processes were shaped into, and named, a theory of emotions by John Dewey (1894). As the saying goes, whoever tells a tale adds a tail, and Dewey seems to have done that because he claimed that James assigned a distinct biological state to each category of emotion (Barrett, 2017a; 2020). However, what James postulated was that each instance of an emotion might be linked to a distinct biological state (Barrett, 2017a; 2020). As instances are different from categories, the tail seems to have become a tale and interpretations led to some confusion.

However, important caveats to the so-called James-Lange proposition soon emerged. By the end of the first quarter of the 20th century, Cannon (1927) pointed out that physiological changes - by being more stable and slower in processing stimuli - occurred independently of behaviours deemed emotional. In a later development of experimental findings derived from visceral phenomena, Cannon, together with Philip Bard (his doctoral student at the time), postulated a theory of emotions based on an uncoupling of the emotional expression, processed by the hypothalamus, and the emotional feeling, processed by the dorsal portion of the thalamus. Cannon (1927) contested the notion that visceral processes underlie emotions stating that: “responses in the viscera seem too uniform to offer a satisfactory means of distinguishing emotions which are very different in subjective quality” (op. cit., p. 110). Importantly, Cannon pointed out that afferent impulses from viscera were not the total picture and that subcortical centres, such as the thalamic complex, were the driving forces of emotional expressions. Of note, Cannon encouraged the notion that subjectivity did play a part in emotions.

In 1962, Stanley Schachter and Jerome E. Singer published a paper entitled “The Cognitive, Social, and Physiological Determinants of Emotional State” relating findings obtained from an experiment conducted at the University of Minnesota as part of a research program based at Columbia University. Findings confirmed the authors’ hypothesis that cognitions may explain physiological arousals and, should they be lacking, i.e., in the absence of an appropriate explanation, the sympathetic activation (physiological arousal) may give rise to a variety of emotional states. Theirs was a two-component model (general arousal and pertaining cognitive interpretations resulting in an emotion) that paved the way for psychological constructionist views. Their conclusions were crucial in spreading the notion that “cognitive factors appear to be indispensable factors in any formulation of emotion” (Schachter & Singer, 1962, p. 398).

Following these theories, there were great efforts in explaining emotions. Most, however, have arisen from central tenets of each theory. Pribram (1970) and Nauta (1971), for instance, built on the ideas of James and Lange on physiological arousal. Specifically, in investigating the role of the prefrontal cortex, Nauta (1971) explored interoceptive information, i.e., that information whose “pre-setting could be thought to establish a temporal sequence of affective reference points serving as ‘navigational markers’ and providing, by their sequential order, at once the general course and the temporal stability of complex goal-directed forms of behaviour” (op. cit., p. 183). Pribram (1970) defined emotions as “neural programs which are engaged when the organism is disequilibrated” (op. cit., p. 43). To Pribram, emotions should be conceived as ‘plans’ engaged to provide for the arousal, or reaction, that the organism experiences when a reason emerges. Both notions, i.e., *interoceptive information* (Nauta, 1971) and *emotions as plans of neural programs engaged in an unbalanced state* (Pribram, 1970) – have grounded subsequent works by Damasio and colleagues (Damasio, 1994; Damasio et al., 1991) who developed the notion of *somatic markers*, i.e., physiological modifications in visceral and musculoskeletal systems that signal the brain to release neurotransmitters, cause changes in somatosensory maps, and modify signalling from the body to somatosensory areas. As much as Pribram (1970) distinguished emotions from feelings, Damasio (1999; 2003) proposed that responses from the body (or “soma” in Greek) and the brain compose an **emotion** while signals in somatosensory regions constitute an individually accrued perception of a **feeling** (Bechara & Damasio, 2005). In the next section, we will turn our attention to what that distinction implies.

**1.4 Emotions: Different Viewpoints**

Current understanding of the brain takes it to be an apparatus of metabolic efficiency that regulates the interfaces of the organism’s internal and external environments in ways that are optimally conducive to its constancy and reproduction (Barrett, 2017a; 2020). This is attributed to its primary function of managing metabolic resources (e.g., enzymes, neurotransmitters, hormones) via their designated cells (neurons and glia) in an efficient way to keep our bodies alive and active and thus ensure the reproduction of the species (Barrett, 2017a; 2020). Doing this requires resources, such as glucose uptake, to be used as long as there is a need, such as reacting to a stimulus perceived as threatening. What follows is that there must be a reason for resources to be directed to one action (to the detriment of another) or to one decision (and not to another). Indeed, the brain, as part and parcel of a biological system, seems to optimize cost functions (Marblestone, Wayne, & Kording, 2016).

Our brain seems to construe meaning in several ways - via internal information (such as the burning of the stomach denouncing a possible inflammatory state or anxiety) and external information (such as the smell of burning that can mean a fire) and its interaction (in the previous examples, burning is a common denominator with differing meanings leading to diverse interpretations). That’s why context – both internal and external – needs to be appraised. This means that each brain needs to interpret information from its internal and external world.  When information comes from external sources (captured by our senses), the brain interprets sensations and performs the same interpretative process with regard to information from internal sources (interoceptively meaning using autonomic, hormonal, visceral and/or immunological signals, Barrett & Simmons, 2015). That is the underlying reason for the concept of ‘burning’ being distinct to designate an interoceptive state, like the burning in one’s stomach, from a sensorial perception, like the burning smell. Both refer to instances that generate conceptual knowledge.

For the perception that we are feeling something to arise, however, it is necessary for the brain to pair the new stimulus (whether internal or external) with some previous information that allows its recognition. This information is prior knowledge, or *priors*. These *priors* form the internal model of the world that each brain builds and that allows the brain to make an inference based on an inverse probability. This means that the same effect/pattern of stimulation can have many different causes. Such inference - starting from what is known to infer probabilities about what this new stimulus means – is an *a posteriori* probability. This way of treating variables by assigning them a value in a probabilistic distribution was first described mathematically by Thomas Bayes in the eighteenth century. Later, this probabilistic distribution came to be known as *inverse probability* - because of the inference made in reverse, i.e., from the effect to the cause - until, in the middle of the twentieth century, it became known as Bayesian statistics (Fienberg, 2006).

Currently we have a set of possible explanations, which have become a theoretical framework - with elaborations still being refined - that manages to account for how experience, in the form of memories, and salience, in the form of attention, compose a scenario of sensory anticipation that allows us to perceive the stimuli present in our context – internal or external - at each moment (Gardner & Gardner, 2023). This is the framework of the Bayesian Brain which we will further discuss below.

The Bayesian Brain – a name among others that designates a predictive process – answers for a theoretical framework for how the brain operates, in neurobiological terms, to process information. In general, the theory presupposes the existence of an internal model of the world embodied in each brain that generates predictions at all times about the causes of what we perceive (Friston, 2003; 2005; Rao & Ballard, 1999). This is accomplished through information we get from our environment, whether external (such as sensory stimuli) or internal (such as visceral information). This framework is based on the idea that the brain makes an inference (Friston, 2003; 2005), that is, we have the result (such as the burning sensation in our gut) but we still don't know what exactly is causing that sensory information. In the example above, the symptom could be physical, such as heartburn, which requires prompt attention to prevent discomfort. It could also indicate a mental state, such as anxiety, which likewise deserves immediate consideration as it directly affects our emotional well-being (Barret & Bar, 2009).

It is through the experiences we collect, in the various environments in which we live and through the succession of experiences with similar results (a burning sensation is not so easily forgotten) that we achieve an internal repertoire (i.e., our internal model of the world) that prepares us to receive information from external sources (such as an upcoming test that may cause some anxiety) or internal sources (such as the restlessness that ensues from a burning gut, which lead us to reason that something is not all right). Depending on the intensity of the stimulus (such as whether the burning is too strong and accompanied by nausea), we make a decision based on information we have (such as the need to look for professional help), and thus, we recalibrate our internal model, generating one that contemplates different possibilities for the same originating source.

It is at this point that the theoretical framework lives up to its designation - that of a probabilistic processing - which receives several denominations according to components and viewpoints (see Table 1 below). Basically, the designations refer to how the brain makes an inference based on an *a posteriori* distribution of probabilities (all the possibilities and their expectations of becoming reality listed above for the example of the burning gut). The designation is justifiable as it presupposes the weighing of probabilities to make inferences. In this framework, the brain recognizes and categorizes information allocating attention and mobilizing resources where action is needed. For a more detailed account involving algorithmic rendition of diverse models of predictive coding, readers are referred to Spratling (2017).

**Table 1**: Denominations, selected examples sources, and theoretical frameworks used to designate probabilistic inference performed by the brain.

|  |  |  |
| --- | --- | --- |
| Denomination | Example sources | Theoretical framework |
| *Bayesian Brain* | Clark, 2016  Doya, 2007  Friston, 2012  Hohwy, 2016  Parr, Pezzulo, & Friston, 2022  Seth & Friston, 2016 | The nomenclature is based on Bayes' model (*a posteriori* probability) to explain the prediction mechanism (i.e., inference) about the possible cause(s) of new information based on *priors*. This process is deemed to update our internal model of the world  . |
| *Predictive Coding* | Clark, 2016  Friston, 2009; 2010  Huang & Rao, 2011  Rao & Ballard, 1999 | The nomenclature deals with how neural responses are encoded (in algorithmic terms conceptualized as data compression) based on prediction error. The denomination accounts for a simplified version of the free-energy minimization scheme (Spratling, 2017). |
| *Predictive Processing* | Clark, 2013, 2016  Friston, 2005; 2019  Hohwy, 2012  Seth, 2014 | The nomenclature designates action instantiated by motor and autonomic reflexes that obey Bayesian filtering. The process implies **active inference** centred on prediction error. |
| *Bayesian Predictive Coding* | Aitchison & Lengyel, 2017 | The nomenclature indicates neural responses represented by both prediction error and Bayesian inference. This latter element represents behaviour. The use of both terms aims at encompassing restrictive uses of the terms in *Predictive Coding* and the *Bayesian Brain.* |

Note that what is described above occurs from the inside (in this case, from the brain) to the outside (operating on the stimulus received) and not the other way around. In this sense, beliefs and explanations are not taken as conscious mental states, but rather as probabilistic distributions that are encoded in our neurons in an unconscious processing (Seth & Friston, 2016). In addition, what is noted is that the representations - here understood as the dynamical pattern used in neural networks that encodes processes such as attention (e.g., to list the visceral stimulus of burning as relevant to be attended); memory (e.g., of the feeling of previous episodes of anxiety); and reasoning (e.g., thinking that if it is an inflammatory condition, it is necessary to call for competent help) - operate based on the information received (by visceral afferent pathways) and determine the execution of a motor action (to be performed by our body), denoting one of the organizational principles of the brain: its hierarchy.

It is important to emphasize that the process of all the operations exemplified above represent a joint probabilistic inference, one in which "best guesses" (such as assuming that the burning sensation may signal a medical condition requiring quick action) about what we experience (the information that reaches us at each moment) act as *priors* for our perception and attention - and vice versa, because when we attend to and perceive something it is because there was an internal model that allowed the identification and relevance of the information received - in a continuous process where attention works to optimize inferences (Clark, 2016; Feldman & Friston, 2010; Hohwy, 2016; 2020).

In summary, our effort in this section so far has been to provide an understanding of a theoretical framework with possible explanations of phenomena evidenced about how the brain processes information to generate perceptions[[3]](#footnote-3) and emotions. This framework, based on predictive processing, offers models for information processing in the brain. Such information, consisting of sensory and interoceptive signals, is the source for emotions. Understanding how this has been conceived in cortical layers may benefit machine learning tasks (Tscshantz et al., 2023).

We adopted the nomenclature of Bayesian Brain, although there are several denominations and possible understandings with different nuances that imply algorithmic adjustments (as explained in Table 1). In general, this understanding regards the brain as an organ capable of an active inferential process that generates predictions based on sensory (proprioceptive), interoceptive and external (exteroceptive) information. This perspective entails conceptualizing the brain as an entity capable of constructing and continuously refining a generative model of the external environment, utilizing probabilistic inference to assign likelihoods to the potential causes underlying incoming data (Friston, 2010; Hobson and Friston, 2012; Picard & Friston, 2014; Friston and Frith, 2015). In this sense, active inference translates into how the brain interprets sensory prediction errors, i.e., actively (Seth & Friston, 2016). Importantly, each brain, when managing resources economically, seeks the most likely explanation for causes/information that may remain inaccessible (Palmer, Seth & Hohwy, 2015). That is why emotions stand a greater chance of being better understood as heuristics in communicating internally generated instances demanding actions.

We know that the understanding of "what" happens in the brain for emotional processes to have full development implies using a scientific model to understand a certain phenomenon, i.e., the *map* of the neural territory that encompasses representations, distributions and transformations. We also know that this understanding differs from the understanding of "how" these emotional processes operate and enable the full functioning of the cognition that underlies the individual's performance. This implies a neural system that specifies implementation i.e., the *territory[[4]](#footnote-4)*. Such understanding permeates, for example, the notion that evidence on neural activation patterns consists of datasets with probabilistic treatment whose nominal values can reflect enormous variation (Barrett, 2017a; 2017b; Reitsema et al., 2021). In other words, it is one thing to map the functional layout (Van Essen et al, 2017) and another to account for how each brain processes information (Krieskorte & Douglas, 2018). Caveat made, we proceed with an effort to pave understanding starting with a theoretical framework for the functional layout.

Based on evolving evidence from sensory systems, such as vision (Peelen, Berlot, & de Lange, 2024; Perrinet, 2020; Rao & Ballard, 1999), we now know that perception occurs via an active inferencing process (involving enaction or embodiment), predictive by nature, where affective valence seems to play a key entry role (Barrett & Bar, 2009; Shenhav, Barrett, & Bar, 2013; Lebrechet et al., 2012) but where more nuanced aspects (such as brightness in relation to visual perception) still require deeper analysis (Jacobson et al., 2024) .

For perception to occur, iteration is required. As stated by Miskovic and collaborators (2016): “perception is more properly conceptualized as a dynamically reverberating loop rather than encapsulated bottom-up and top-down streams.” (op. cit, 2016, p. 41). In this sense, context is constantly informing predictions and neurons operate differently according to the task involved (Hackel et al., 2016). That is why a view from the brain and the mind, might unite what history has biasedly separated. Clear-cut definitions between perceptual, cognitive and emotional processes stem from more theoretically driven experimental studies about how brain and mind operate that have been criticised (Hackel et al., 2016; Miskovic et al., 2016). However, earlier notions of affect’s role in perception, steered by Wundt in the late nineteenth century (as stated in the first section), seem to have found traction in recent studies. These propose that neural substrates that code for affective dimensions also work for physical properties of the stimuli (Miskovic et al., 2016). In summary, brain states that reflect subjective experiences are observable and overlap with those that have been associated with cognition (Todd et al., 2019) and also with perception, action and motivation (Pessoa, 2022) making emotions a whole-brain state (Malezieux, Klein, & Gogolla, 2023).

To elucidate the emergence of dynamic loops from the processing of information streams, we now present a concise and foundational overview of information processing mechanisms. This may be considered elementary by individuals possessing prior knowledge of brain mechanisms. To those, please feel free to skip to the next paragraph. Inside the brain, primary areas, or areas lower in the hierarchy, are those that mostly receive sensory afferents while associative areas, or higher areas in the hierarchy, mostly process multiple modalities related to the same stimulus. In neurobiological terms, the underlying reason resides in the direction of the connections between neurons – the nerve cells responsible for signalling in the nervous system.  This signalling happens in the brain when electrical signals (or action potentials) propagate between a set (or network) of neurons that connect (or synapse). For information to be communicated, primary or afferent sensory neurons carry data to the Central Nervous System (CNS). In order for data to inform actions, information is communicated from the CNS to motor organs by efferent or motor neurons.

Hence, diverse patterns of interconnection between neurons allow the transmission of different information – for perception or action - according to the anatomy or functioning of the system in demand. In general, ascending projections of primary neurons to the CNS are called *feedforward*, as their axons project from lower to higher areas of the hierarchy. Neurons whose axons have descending projections, usually modulating motor commands, are called *feedback* (Amaral, 2023; Berezovskii, Nassi, & Born, 2011).

A growing research body, based on experimental studies in differing mammalians, makes evident that feedback connections take more time to mature than feedforward ones (Berezovskii, Nassi, & Born, 2011). While hypotheses of cost functions that change across development and areas in the brain lend more recent support to the idea (Marblestone, Wayne, & Kording, 2016), a prolonged maturational process supports the notion of feedback as predictive because it relies on a mature feedforward circuitry possibly aligned with visual experience (Berezovskii, Nassi, & Born, 2011). The same kind of support justifies a certain vulnerability that nature and nurture selectively impose on feedback connections generating auditory hallucinations, and visual illusions, such as that regarding the size-weight distortion or the hollow-mask illusion (Berezovskii, Nassi, & Born, 2011).

In the cerebral cortex, the portion closest to the surface - the neocortex to many, also referred to as isocortex by some (e.g., Pessoa, 2022) - contains layers and columns on its extensive surface that were made possible evolutionarily by the number of grooves and gyres present in mammalian brains (see Figure 6). This had a relevant consequence: there was an increase in computational efficiency between afferents and efferents in each of the six cortical layers composing the brain’s gray matter. The neural organization regarding brain functioning occurs in columns, and each column - which runs through the six cortical layers - makes up a computational module that aggregates neurons with a similar functioning pattern thus forming a specific and local processing network. To get an idea of the specificity of brain functioning, the primary somatosensory cortex, which processes the sensation of touch in fingers, encompasses many cortical columns and contains four somatotopic maps, or neural representations, of the skin that are connected in a distinct way. Each map is responsible for processing specific information, and is responsible for a certain function (Amaral, 2023).

Diagrama

O conteúdo gerado por IA pode estar incorreto.

**Figure 6**: Illustration of a column (left) and types of cells (right) illustrative of the composition of the six-layered mammalian cortex from Jones (2000) . [pending authorization]

In summary, each sensory modality maps onto a specific cortical area with a dedicated set of cortical columns, whose processing occurs through a series of circuits, with increasing complexity, in areas of unimodal association to later converge to areas of multimodal association. Of note, specialization becomes a central tenet for the brain’s efficiency. From what has been evidenced so far, patterns of information flow differently across brain areas to provide an efficient solution to computational problems (Marblestone, Wayne, & Kording, 2016).

Essentially, there are two-way processes that seem to be happening in the brain regarding stimuli. For stimuli captured by the sensory organs to become perception, there needs to be certain regularities. Those are obtained via statistical computations that are performed from birth in sensory-motor patterns. These patterns, formed across time and in multiple interactions, constitute internal representations that are stored in the brain. They become prior knowledge. When stimulated by incoming information, the brain predicts what this new information might be based on using priors. Priors act like signposts that allow for the recognition of stimuli to be processed by dedicated neural networks (Barrett & Bar, 2009). Regarding the incoming stimuli that may be proprioceptive, interoceptive and exteroceptive, which are being continuously captured, the path will feature: (1) stimulation of receptor cells in specific parts of the body; (2) transformation of the captured stimulus into neural representation via fragmentation of the stimulus into components encoded by dedicated projection neurons; (3) serial and parallel upstream uptake of neural representations; (4) modification and regulation of the signals received via feedforward and feedback projections; (5) integration of the representation of the input (such as an object, person or scene) by neural networks; (6) selection of abstract features based on detailed input; (7) salience of features according to temporal/spatial importance (Gardner & Gardner, 2023). Interestingly, such features form the internal representations that compose sensory-motor patterns. They have been referred to as affective (Barrett & Bar, 2009) as they affect how we perceive something. Affective value – the percept’s characteristic of being experienced as good or bad - imbues perception, even for faces in the absence of valence configurations (Siegel et al., 2018). That absence is noteworthy.

Compounding and converging evidence from sensorial afferents seems to indicate that sensory afferents that carry affective-valence information are processed by different nerve fibres from those that carry valence-independent information (for a review, see Kryklywy et al., 2020). According to this viewpoint, abstract cognitive states, i.e., emotions, are the accrued result of “the perceptual experience of specific exteroceptive sensation of objectively valence objects” (Kryklywy et al., 2020, p. 926). In the end, we have perception as an affective, mental/cognitive construction that uses environmental – either internal or external – stimuli. Not the other way around.

Integral to this view is the issue of valence-coding. In tandem with psychological viewpoints, current neuroscientific evidence points towards the relevance of valence for survival mechanisms that maximizes resources and minimizes threats (Kryklywy et al., 2020). Valence-coded information, those denoting affective states, is crucial not because they are endpoints in a dimension but rather because they operate as superordinates, i.e., they organize one’s behaviour and experience (Kryklywy et al., 2020). A recent review (Malezieux, Klein, & Gogolla, 2023) highlights that neural underpinnings in the processing of emotions should be uppermost in future studies. This is because semantic hyperordination, such as the broad categorization of emotions, fails to account for distinct aspects that require more precise nomenclature. Greater specificity would facilitate deeper understanding of features, interpretations, contextual evaluation, and adaptive expression (Malezieux, Klein, & Gogolla, 2023). An essential part of how such distinctiveness might unravel lies with understanding the performance of each set of neurons, that is, the character of the transmitted and sustained signal.

Most of the neurons that relay sensory information in the thalamic nuclei receive excitatory signals in a convergent manner – as in many to one. In information integration, relay hubs combine inhibitory signals at the layers and columns before passing the information on to multimodal association areas. This means that we indeed have an efficient encoding by sensory neurons; efficient because it is able to reduce redundancy without losing the essence (Gardner & Gardner, 2023). This process occurs laterally, that is, relay neurons within the same layer transmit excitatory signals, forming recurrent networks capable of amplifying the signals of the stimulus being processed. Meanwhile, interneurons are designed in such a way as to contain – via an inhibitory signal - the amount of sensory information that is relayed to the upper layers. This inhibition is important because it reduces irrelevant information by focusing attention on neural pathways that effectively lead to the achievement of the task (Gardner & Gardner, 2023). It is throughout the processing, via excitation and inhibition, in a parallel and serial way, that the spatial arrangement, i.e., the dedicated and specific processing of the stimulus components, lays the groundwork for relevant features that the stimulus might have for behaviour (Gardner & Gardner, 2023).

It is important to note that each of the primary sensory areas is endowed with more feedback axons than feedforward ones (Gardner & Gardner, 2023). This means that higher areas modulate sensory receptor responses. If a form of comparison could capture—however faintly—this processing from sensation to perception, the process would not be like a standard relay race where the baton is passed from one [neuronal assembly] to another till the finish line. Rather, what we have is a process of constant restructuring, as if in order to add more value and produce with more quality, the runners [connecting neurons] had to redistribute roles and reorganize to speed up the process, as in “The Amazing Race”[[5]](#footnote-5). In this way, perceptions and affective states are subjective because they depend on the capacity for continuous reorganization of each brain.

Circling back to the brain as an optimizer of cost functions, evolution seems to have forged a way for unsupervised learning to conjure up relatively limited input with the right output, that is, an individually weighted, adequate behaviour that sometimes demands recruitment of selective variables, over long timescales for coherent action (Marblestone, Wayne, & Kording, 2016). This seems possible via an internal bootstrapping process whereby one template is used to generate different goal-targets that attend to skill development (Marblestone, Wayne, & Kording, 2016). Whether it may be taken today as amodal due to an evolutionary refinement of specific sensory events in our interaction with the outside world (Kryklywy et al., 2020) still requires more concerted efforts. However, the similarities between what the brain does - and neuroscience may reveal - and machine learning operations regarding cost functions should not be equated (Bengio et al., 2009). There is much more that the brain does regarding supervised, unsupervised and reinforcement learning to account for the adaptability and suitability of core processes, like facial movements, to convey individual needs within their ecology.

**1.5 Interfaces between Neuroscience and Emotions: What have we learned**

In science that caters for how humans evolve and develop, sound reasoning and research should draw the line between where consensus and dissensus lie and show assertions through documentation (Bornstein & Lamb, 2015). On the topic of emotions, that documentation finds in neuroscience a major throughfare as it confers legitimacy and explanatory depth to neural correlates (Busso & Pollack, 2015), i.e., biological tissue, of emotional substrates.

The biological tissue in the brain shows excitability when neurons fire. Their spikes may be registered in terms of (a) frequency, measured in hertz (Hz) accounting for how many times neurons fire, i.e., metabolize oxygen and glucose, at each time interval; (b) amplitude, measured in microvolts (μV) accounting for how many different sets of neurons fire at each frequency; and (c) potency, obtained by the amplitude squared (measured in microvolts or mV, where  1 mV = 1,000 µV) accounting for the relative size of the neuronal set firing at each frequency. That is how measurements are obtained. Let us turn to what those measurements are tracking.

As described in the previous section, neurons in cortical layers interact, via association and projection fibres, to enable network formation from which ensues mental states and activity. The highly dynamic nature of such neural networks produces oscillations that can be recorded via electroencephalography (EEG). Circuits within cortical columns and those encompassing different brain regions characterize complex neural networks that enable sensory, motor, and cognitive functions with distinct oscillatory frequencies and amplitudes. Different sets of neurons synapsing into a network can represent the same category of information, such as an emotion category, creating a spatiotemporal pattern. That is called “degeneracy” (Barrett, 2017 after Edelman & Gally, 2001) where different forms come to shape the same pattern. In general, local networks operate in fast frequencies while more diffuse networks, covering large distances, present slower frequencies justified by the delay implicated in their distancing (Aboitiz, 2024). The same has been noted in relation to emotionally salient facial features (Symons et al., 2016).

Understanding the “what” and the “how” of neuronal coding measurements opens up a window into the mechanisms underlying energy patterns, or the amplitude of neuronal modulations taken at many rates simultaneously (Buzsáki, Logothetis, & Singer, 2013; Doelling & Assaneo, 2021). Brain waves, or oscillations, encode rhythms entrained in patterns. When many neurons fire at rhythmic patterns, they encode oscillating networks – triggered by internal and/or external ongoing activity - that together generate perception (Calderano et al., 2014). Of note, research has highlighted how ongoing, or internally generated, oscillations are capable of modulating responses that shape perceptual, motoric, and cognitive processes (Calderano et al., 2014; Thut, Miniussi, & Gross, 2012) underlying emotions. Oscillatory phenomena have reshaped our understanding of some aspects of emotions, such as memory (Headley & Paré, 2013), regulation (Popov et al., 2012) and expression (Symons et al., 2016).

The takeaway for emotion research is twofold: (i) Degeneracy, which refers to the phenomenon whereby similar emotional experiences can arise from distinct spatiotemporal neural configurations, challenges the notion that facial expressions are fixed, universally recognizable indicators of emotion, as this view is not substantiated by contemporary neuroscientific evidence (Barrett, 2017b); (ii) Ongoing neural activity, encompassing more than discrete stimuli derived from sensory or interoceptive inputs, undermines the stimulus–response paradigm traditionally used to conceptualize emotions. Rather than being elicited by isolated stimuli, emotional experiences emerge from the continuous engagement of neural systems with contextual information (Barrett, 2017b). What follows is that there does not seem to be a dedicated brain region or system for how we feel, think or act (Barrett, 2017b). In other words, emotion, cognition and perception share neuronal networks devoted to entraining patterns that make efficient use of the metabolic resources that the brain manages to keep us alive, developing and reproducing[[6]](#footnote-6), i.e., keep us active in the world (Barrett, 2017b).

Pezzulo, Parr and Friston (2021) posit that, to achieve adaptive control, organisms must reduce the number of possible states to align with the specific demands of their environmental context and ensure survival. Whatever lies outside the boundaries of such states constitutes a surprise, i.e., anything that may threaten adaptative survival computed via sampled sensory data combined with their internally generated predictions (Barret & Simmon, 2015; Barrett, 2017b; Holmes & Nolte, 2019). Thus, the brain - in experiencing different bodily states when combining internal prediction with incoming evidence - is constantly comparing experiences with expectations. Ensuing perceptions may contain prediction errors which are, in typical conditions, expertly used to modify future behaviour as an emotional display (Platt & Spelke, 2009). A caveat is necessary, though, and that has to do with the free-energy principle.

According to the free-energy principle, living organisms continuously engage in inferential processes to counteract their inherent tendency toward entropy (Friston, 2011). This is achieved through the construction and iterative refinement of an internal generative model that represents the structure of the external world. Through sustained interaction with the environment and the accumulation of sensory evidence, this model is progressively updated. Consequently, each organism, by enacting behaviors through its bodily interface with the environment, contributes to the emergence of an adaptive phenotype optimally attuned to its ecological niche (Friston, 2011). Reaching optimal levels of adaptation means reducing the amount of free-energy; therefore, embodiment becomes an integral part of the statistical process that allows for an internal model of the agent in that environment (Friston, 2011). That means, for the brain to generate inferences, the context is important (Friston, 2011).

When the context holds more novelty potential, expectation increases, i.e., the more surprises the context offers, the less accurate the inference is likely to be (Pierkaski, 2021). In other words, inferences with low accuracy upgrade our level of surprise (or novelty expectation) and make our senses more open to new information, i.e., senses tend to get more tuned in, more precise, as they become more necessary to update one’s own generative model of the world (Pierkaski, 2021). On the other hand, inferences with high accuracy are used to represent the causal structure of the world and of one’s own bodily states (Pierkaski, 2021). This means that they become heuristics, i.e., perceptual shortcuts, that are engaged frequently to make sense of information. This may engender a vicious circle that tends to fixate a mechanism – of perceptual inference – that should never be fixed, to allow one to maintain a healthy perspective and stance in the world. For example, when you see a snake, even living in an urban environment where snakes are not prevalent, you may get goose bumps, thus you infer – either consciously or not – that you must be afraid. Consider now an encounter with a piece of a rope; you may take it for a snake, get goose bumps and feel afraid. In this case, if you use a mechanism that allows you to balance your former experience (of seeing a snake) against the latter experience (with a rope), you will confer a value that renders each experience in a precise way. The mechanism would mean that, when your brain notices that the information looks the same but behaves differently (in the example above, realizing that a rope does not slither), it would be performing a weighted precision balancing act (Clark, 2013; 2017; Seth, 2014). In this act, the brain generates an error signal (e.g., realizing the information is a rope and not a snake) resulting in a prediction error (Hohwy, 2012; Seth, 2014). This error is sent forward and sideways via distinct networks, i.e., in serial and parallel processing pathways, that will probably improve the perception you have of future input that looks like a snake, but that is most likely a rope and thus will not give cause to a feeling of fear. Learning has taken place and the error that has been previously made (e.g., of thinking that the rope you saw was a snake) is taken as a model of neural responsivity (Heilbron & Chait, 2018). In this example, the fixed mechanism would be taking every rope encountered as being a snake. Most likely, that would imply the agent is not coming close to, nor interacting with, the environment where the object (rope, most probably) lies. That would perpetuate the assumption and reinforce the fear that could be reasonably justified (in case of a snake) but not in case of a rope, especially considering the urban context where the agent is. The takeaway message here is to understand that perceptions are inferred based on constructed sensorial and interoceptive stimuli placed in a context so that interaction with that input happens as a constant, ongoing activity, forging ways of forming knowledge representations in the brain.

Let us turn now to how formed knowledge is represented in the brain through our motor cortices or via action. To think of action related to emotions and sensations, one must consider that the body has to be factored into the equation, i.e., taking some level of action - either using one’s own bodily parts or imagining their use when physical action is not feasible. This account challenges the classical views of cognitive psychology where mental representations get entrained irrespective of their sensorimotor afferences (Macrine & Fugate, 2022).

Sometime before and throughout the 20th century, philosophers like Gilbert Ryle and Maurice Merleau-Ponty together with psychologists like John Dewey, William James, and James Gibson severed this notion by pointing out that mental and physical worlds merge into bodily perceptions of information to generate knowledge and states (Macrine & Fugate, 2022). In the present century, Clark (2013, 2016, 2019) emphasizes the role that active inference plays for the brain to generate predictions about the world. This is seen when one’s own senses, i.e., bodily mechanisms, are actively engaged in interactions with contextual information or input.

Historically, the idea that body and mind converse in transactional, recursive patterns emerged in this century based on the tenets of (i) embodied realism; where the body grounds mental knowledge; (ii) sensorimotor simulations; where perceptual symbols arise based on sensorial afferents; and (iii) situated cognition, where context matters for the representation of abstract, mental concepts (Barrett & Lindquist, 2008). Therefore, all forms of cognition and emotion come from an embodiment where experience becomes the product of model-based predictions that are flexible, i.e., not fixed or immutable, as the same elements may yield different experiences over time (Barrett, 2017b; Clark, Friston & Wilkinson, 2019).

As one’s knowledge depends on structures shaped by experiences that one has, and as experiences may vary according to culture, norms, and language one adopts (Macrine & Fugate, 2022), it follows that embodiment offers the possibility of forging patterns that serve as heuristics when interacting with the world (Clark, 2016). This may be taken, at the surface level, to explain why smiles are taken as symbols for happiness at a rate that is more often than chance but, at a deeper level, it makes more plausible the reasoning that smiling may communicate different instances of emotion, such as surprise and even sadness, or a combination of both, when cultural values and constraints are factored in (Barrett et al., 2019).

Taken together, perception from sensory and somatosensory cortices together with action involving motor cortices are entrained along multiple space and time scales and interactions shaping how the brain simulates emotions and forms conceptual knowledge (Barrett, 2017a; 2017b). In this framework, and as we have seen before, language plays a constitutive role in emotional experience (Gendron et al., 2012; Lindquist & Gendron, 2013; Satpute & Lindquist, 2021). This derives from the notion that language drives the acquisition of categorical patterns (see the Chapter on Language by Monte-Serrat & Cattani in this book) formed via automatic (largely sensorial and visceral), behavioural (how we display our emotions in the world) and experiential (how action enables experiencing emotions) information that forms the building blocks of individualized conceptual emotional knowledge (Barrett, 2017a; 2017b).

From the neuroscientific input we have laid out here, emotions may be taken as the compilation of affect – that are innate and biologically driven – and conceptual knowledge – formed through inferences – forming patterns, or categories, that imply changes in states for the emergence of a behaviour, i.e., instances of emotions that we perceive in others, via facial configurations. These configurations are made or constructed internally via inferences which make them highly flexible (Barrett, 2017a). Language is a means – a culturally shaped cognitive tool – that enables one to categorize emotions (Barrett et al., 2001) parsing them into specifics that allow us to differentiate between them (Tugade, Fredrickson, & Feldmann Barrett, 2004). In other words, the highly flexible system that configures emotions is the product of affects (innate, biological and basic) and conceptual knowledge (acquired from prior experience and language). Given its exposure to diverse cultural contexts and individual-specific factors, this adaptive system exhibits considerable variability, thereby challenging the notion of emotions as invariant signatures and positioning variation as a fundamental characteristic (Barrett, 2017a; 2017b).

Further, because emotions are formed by conceptual knowledge derived from probabilistic estimations, they cannot be construed as reactions. Instead, they become simulations, here taken to be constructed internal representations of long-term prospections (i.e., predictions) that are formed by deterministic short-term processes (i.e., inferences) (Barrett, 2017a; 2017b; Brown & Brune, 2012). As Barrett (2017b) posits, there is a “lesson here, for the science of emotion, [in] that the brain does not process individual stimuli—it processes events across temporal windows. Emotion perception is event perception, not object perception” (op. cit, p. 12).

**1.6 The Search for Biomarkers: Breakthroughs and Thresholds**

As previously seen, understanding emotions requires a dive into the seat of its processing - the brain - here understood as the set of structures and networks, in and among cortical and subcortical regions, that underlie behaviours and that are part of the central nervous system. By dive we mean an incursion into how processes develop to underlie networks and enable neural systems. To address processes, we have opted in the present work to restrict explanations to encompassing terms, like cortical and subcortical regions, and have used underlying structures, like the amygdala – which, incidentally, should receive a more thorough treatment given its status as a complex of distinct nuclei (Freese & Amaral, 2009) - only when strictly necessary. This approach was developed with careful consideration of the intended audience bearing in mind the implications of an accurate neurobiological analysis of structures and their associated terminology. Another choice concerns theories that inform how data is treated.

As previously discussed, the field of emotion research is characterized by a proliferation of theoretical perspectives, with limited integrative efforts aimed at synthesizing these viewpoints to advance the discipline. Notable exceptions exist (e.g., Barrett et al., 2019), though some attempts remain insufficiently informative (e.g., Dukes et al., 2021). Consequently, this section concentrates on examining three prominent theoretical frameworks—affective feelings, homeostatic feelings, and constructed emotions—highlighting proponents and their foundational contributions to the understanding of emotional phenomena.

Affective neuroscience has emerged as a discipline to address the need for a comprehensive examination of emotional information processing and the mechanisms underlying expressive behavior (Davidson & Sutton, 1995). At the time, it was seen as a field that merited a separate treatment from cognitive neuroscience, perhaps mirroring the siloed treatment that emotions received. Back then, it was attributed mainly to a persistent view of independent and/or nonrelated treatment of subcortical substrates (e.g., the amygdala) as an apt way to investigate emotional phenomena. Also, the siloed treatment was observed in the consideration given to processing sites (e.g., the prefrontal cortex) as sets of “pure” cognitive phenomena (Davidson, 2000). Fortunately, this siloed route soon took a turn to convey research that addressed the pertaining and inextricably related cognitive processing of emotional substrates (Davidson, 2000).

A more contemporary view addresses emotional science based on neurobiological substrates that consider the dangers of regarding affect and cognition as (i) present in distinct parts of the body; (ii) independent from each other; (iii) in the realm of siloed perspectives; (iv) species-specific; and (v) comprising of totally conscious phenomena (Davidson, 2003). That said, and perhaps to show how emotional science progresses much like the nonlinear way that the brain operates (Vignesh, He, & Banerjee, 2024), we proceed in this section with an examination of the breakthroughs provided by three key exponents of affective neuroscience. After each contribution, we summarize their main ideas and weave them into the great tapestry of understanding that such breakthroughs represented. Lastly, we look at the most recent findings that unfold the possibilities of research in the field of emotions.

*Some Breakthroughs*

Similarly to most specific fields of study, affective neuroscience has an exponent regarded by many as its 'father' (Spektrum der Wissenschaft, 2012), in the late psychobiologist and neuroscientist Jaak Panksepp, a first proponent of the [then] emerging field of studies (Panksepp, 2003, p. 2), who got some fame for tickling rats. Through his experimental work, Panksepp discovered that rats emitted ultrasonic vocalizations—inaudible to the human ear—when subjected to tickling, indicating a form of laughter. Utilizing animal models, Panksepp established the biological underpinnings of affective neuroscience by identifying a set of primary affective systems, which he conceptualized as fundamental categories of emotional consciousness that influence behaviour. These *affective feelings* involve distinct neurobiological processes - different from emotional and motivational feelings - driving organisms towards cognitive choices, as in what to eat when hungry (Panksepp, 2003).

To Panksepp, affective systems would be localized subcortically and bear an evolutionary - therefore homologous to other mammals - root as they satisfied basic needs, either biological, social, or concerning safety (Panksepp, 1998; Toronchuk & Ellis, 2013). According to Panksepp, affective systems have two valences: positive and negative. Those of positive valence provide some kind of reward to our brain and the negative ones are aggressive to our brain. Among the rewarding systems we have (1) seeking, generating enthusiasm; (2) play, generating fun and joy; (3) care, generating parental affective vigilance and nurturing; and (4) lust, linked to sexual desire itself and relevant for the reproduction of the species, and also reproductive desire, focused on maternal behaviour related to the desire to protect offspring. Among the negative valence systems, Panksepp identified (5) fear, generating anxiety; (6) rage, generating anger; and (7) panic, generating loneliness and sadness ensuing from separation and the unmet need to be cared for (Panksepp, 1998; Panksepp & Wyatt, 2011).

Panksepp (2001) stressed the importance of understanding the basic emotional drivers, i.e., *affective feelings*, humans share with other animals even considering how their expression might be different in the face of our larger repertoire of cognitive abilities. He defended his position based on the similarity of a neurobiological makeup — emotions are in the brain — common across species. Panksepp emphasized the distinction between evolutionarily older affective states—mediated by subcortical and limbic structures and governed by action–perception mechanisms—and more recently evolved, cognitively mediated emotional processes, which involve cortical regions and operate through perception–action dynamics. It is at this point that another neuroscientist, Antonio R. Damasio, stands out and joins Panksepp in teasing apart ideas that often received a unique label. Damasio did this by studying patients whose emotions displayed incongruent patterns.

Damasio, as seen in Section 1.1, makes a clear distinction between emotions, taken as reactions and belonging to the body, and feelings. In his own words (Damasio, 2001) Damasio states:

“An emotion, be it happiness or sadness, embarrassment or pride, is a patterned collection of chemical and neural responses that is produced by the brain when it detects the presence of an emotionally competent stimulus — an object or situation, for example. The processing of the stimulus may be conscious, but it need not be, as the responses are engendered automatically. Emotional responses are a mode of reaction of brains that are prepared by evolution to respond to certain classes of objects and events with certain repertoires of action. Eventually, the brain associates other objects and events that occur in individual experience with those that are innately set to cause emotions, so that another set of emotionally competent stimuli arises. (...) Thus, emotions are not subjective, private, elusive or undefinable”.

To Damasio, emotions taken as action plans would implicate a three-stage processing of (1) what external or internal stimulus triggers the system; (2) what neural correlates put the plan into action; and (3) the action itself (Damasio, 2011). In assuming emotions as an evolutionary trait, i.e., part of a species’ genome, the processing described above would eventually lead to a perception. Following in Darwin’s steps, Damasio (2011) posits that emotions as action programs form a set – like the set of emotions that conceptualize fear, disgust, sadness, joy, anger, and surprise – with enthusiasm and its counterpart (discouragement) regarded as simpler programs. Social emotions, which include embarrassment, shame, guilt, contempt, compassion, and admiration, would be part of a more complex set (Damasio, 2011).

In 2003, Damasio put forward the notion that a feeling would be “an idea of a certain aspect of the body, its interior, in certain circumstances” (op. cit., p.88). Such ideas, in Damasio’s conception, would represent the message contained in internal maps of the body-sensing regions. Of note, feelings would not accrue from actual bodily states, but rather from their representations in the brain (Damasio, 2003).

A central tenet for Damasio is that emotions are distinct from feelings of emotions – a distinction that he proclaimed had been intended originally by William James, but which found confusion in the elaboration of his thinking (Damasio, 2011). Thus, to Damasio (2011), “feelings of emotions are the perceptions of the action program that constitutes an emotion as it unfolds together with the salient representation of the causative object and with thoughts related to the situation” (op. cit, 2011). Feelings would be elaborations by the mind based on emotions, i.e., the mental representation of our bodily state identified through reactions to stimuli present in the internal or external environment (Damasio, 2011; 2021).

In Damasio’s theory, emotions are automatic and feelings, being their consequence, get conditioned as they pass through the filter of our cognition and may have interference from several other neural and nonneural substrates (Damasio & Damasio, 2023). Hence, feelings derive from reactions (emotions) understood as a response to contextual demands, whether internal or external. Further, feelings are modulated by mechanisms that take place in the CNS but involve mental elaboration that integrate stimuli (of an afferent, sensory nature) and responses (of an efferent, motor nature). Thus, feelings emerge as critical mental constructions that imply some level of consciousness in order to be elaborated. Consciousness means our feelings are rooted in the body, which possesses an inherent biological intelligence. This intelligence maintains internal balance, allowing us to function effectively in our environment (Damasio, 2021). In other words, the emergence of consciousness from perceptions and thoughts in an individual organism means that experiences are registered as mental events. This happens due to feelings that regulate our living needs (e.g., hunger, thirst, pain) and metabolic expenditure (e.g., breathing and thermoregulation), and are thus regarded as *homeostatic feelings* (Damasio, 2021; Damasio & Damasio, 2023).

The importance of homeostatic feelings lies in setting a cornerstone of biological evolution as they provided organisms with consciousness which, in turn, gave organisms knowledge about how life could be regulated (Damasio & Damasio, 2023). By allowing the mental inspection of consciousness over life, feelings as sentinels operate under interactive mechanisms, of (i) interoceptive elements regulated by the nervous system, and (ii) viscera and related chemical elements of a non-neural nature, in an ongoing fashion (Damasio & Damasio, 2023).Therefore, feelings as hybrid constructs with organic origin (from viscera) and neural processing (by brain-dedicated networks) would remain dependent on interoception, since they require a notion of what happens within the organism through sensory and mental inspection. It is such inspection that would provide information of a positive or negative valence and of different gradients (Damasio, 2021).

To recap, in Damasio’s framework, feelings are disturbances that exist for a purpose: they are sentinels of our internal states. Therefore, they demand attention and involvement and depend on a physiology that is elaborated, with gradients of quantity and quality that bring imprecision. Feelings become conscious by way of interactive perceptions. This means that they are (i) hybrids; they manifest themselves in the body and depend on mental work; (ii) internalized; they are inherent to oneself, so much so that the reactions and consequences they trigger become individually perceived; and (iii) emotionally modulated; they affect mood states in a positive way (with enthusiasm) or in a negative way (with discouragement), which, in turn, affects motivation. Homeostasis confers biological parameters that impact perception based on emotions. In this construction, perception means inferring mental states based on experiences of the tripartite state that organisms may hold: (i) internal, represented by interoception; (ii) of one’s own body in a physical environment, represented by proprioception; and (iii) of the environment, represented by exteroception. Therefore, emotions are understood as *involuntary internal responses* initiated by certain perceptions that contribute to maintaining homeostasis. These responses generate feelings, which are considered mental phenomena primarily associated with bodily states and secondarily with emotional states, and are continuously influenced by the central nervous system.

Lisa Feldman Barrett also makes a very strong point about how the brain is involved in emotions but not by following the same line of experimentation and reasoning that Damasio has done. To Barrett (2016, 2020), - whose initial efforts relied on trying to reproduce earlier findings, compiling evidence and critically reviewing the literature - brains run bodily systems and aim at metabolic efficiency. This means being frugal in spending resources, i.e., the body is efficient not when dynamically balancing needs (homeostasis), but rather when predicting where needs will be felt (allostasis). Thus, for efficiency, the brain regulates the body based on signals that the body sends to the brain. Such signals are the information that sensory organs (e.g., eyes, ears, nose) and internal organs (e.g., viscera) send to the brain via afferent pathways. Upon receiving such signals, the brain processes them to produce a general summary, one that provides affects (or moods). These are effective results of how metabolic states are experienced.

However, to interpret what such results (or internal states) mean, the brain “backtracks” to infer the most probable cause for the results (or affects). As we have stated before (in Section 1.1), Barrett calls this a reverse inference (Barrett, 2017a). In inferring the cause, the brain attributes mental states that seek to identify, explain and predict behaviours. Thus, emotions are taken as generated by the combination of affect, that is a transient alteration in a neurophysiological state, combined with the knowledge that we obtain in extracting information based on a distributional computation of their value. In this perspective, Barrett (2006; 2012; 2016; 2017a; 2017b; 2020) posits that emotions (i) emerge as mental states that make use of biological substrates also employed in cognitive and perceptual instances; (ii) are formed according to the situation one is in; varying individually depending on the context and demands of each situation; and (iii) generate a highly flexible system that encompasses variation due to the diversity of cultures, languages, and expressions.

In 2015, Barrett & Simmons (2015) proposed the Embodied Predictive Interoception Coding (EPIC) model to show how active inference (see Section 1.1) gets implemented in interoception[[7]](#footnote-7). According to the EPIC model, the cingulate cortex, the posterior ventral medial prefrontal cortex, the posterior orbitofrontal cortex, and the most ventral portions of the anterior insula are jointly responsible for the process of allostasis. Further, Barrett & Simmons (2015) posit visceromotor cortices as drivers of active interoceptive perception. They also diverted the focus for consciousness or emotional awareness away from the anterior insula by redirecting concerted efforts into understanding how interoceptive perceptions get shaped via multiple pathways in a cortical interoceptive network and via ascending pathways in a finding that points towards degeneracy (see Section 1.2).

In summary, Barrett has largely evidenced that individual brains construct their concept of emotions over time as categories of meaning that enable plans for action to attain preferred states. Such states are expressed in behaviours, like facial configurations. In so doing, emotions may hold different definitions as people and places vary. Therefore, *emotions are constructed* by each individual according to the inferences made relative to the internal, such as sensorial and interoceptive input, and external information, such as culture and language, available to them at different instances represented in diverse temporal patterns[[8]](#footnote-8) (Singh et al., 2021). In that framework, variation becomes the norm. And to realize and truly invest in this notion, the field would do well to acknowledge that:

“It will never be possible to measure an emotion by merely measuring facial muscle movements, changes in autonomic nervous system signals, or neural firing within the periaqueductal gray or the amygdala. To understand the nature of emotion, we must also model the brain systems that are necessary for making meaning of physical changes in the body and in the world.” (Barrett, 2017b, p. 16).

*Interim Summary*

Thus far we have offered a closer examination of three exponents in affective neuroscience and their breakthroughs. Panksepp and Damasio provided experimental basis for distinguishing between affects, emotions, and feelings while Barrett’s concerted efforts went towards evidencing how emotions are constructed. Departing from the reference viewpoints established in Section 1, we offered in this section exemplary trajectories of distinct ideological, theoretical and experimental fronts that were instrumental in advancing our understanding of what emotions are. With their work and those of many others, the field presently recognizes that emotions demand a more complex, refined thinking that should encompass body and mental states represented by neural substrates. To move the field forward, their concerted efforts involved a more intensive and directed investigation of neural substrates - the topic we explore next.

The notion that cognition and emotional feelings are inextricably tied (Damasio, 2003) was a breakthrough that needs acknowledging as it was taken as radical when initially proposed (Panksepp, 2003). Of note, Damasio’s centred his work on the distinction between emotions and feelings. This has been consistently evidenced through an emphasis on the externally consequential nature of the first and the privately experienced consequences that configure the latter (Dolan, 2003). Damasio has made significant and influential contributions to the identification of brain regions associated with emotional processing. Starting with the somatosensory cortex (Damasio, 1994), Damasio moved on to incorporate the hypothalamus, basal forebrain and brainstem (Damasio, 2001) and insular and cingulate cortices (Damasio, 2001; 2003), regarding subcortical regions as the first site for neural representations of feeling later repeated in cortical regions (Damasio, Damasio, & Tranel, 2013).

Another important breakthrough was Panksepp’s notion that affects guide our actions and perceptions, being precursors and propellers of other states. As stated in 2001, “our feelings still remain like centres of gravity around which our cognitive apparatus tends to revolve—unless, of course, all our needs are satisfied.” (Panksepp, 2001, p. 139). In relation to neural substrates, Panksepp postulated that affective feelings originate from lower regions (subcortical areas), the processing site of emotional responses (Panksepp, 2003). He was also adamant in drawing attention to the highly inflated role attributed to the amygdala in recent years (Panksepp, 2003). That position, albeit not fully in synch, has also been observed by Gendron and Barret (2017), in positioning the amygdala as engaged in fear poses (rather than in pure fear) and also in positive stimuli and novelty. This extended an earlier notion of the amygdala as a gating, modulatory substrate of emotional input, like facial expressions (Morris et al., 1998).

Barrett's work represents a significant advancement in challenging established beliefs about emotions and examining their implications for the identification of biomarkers. In an authoritative, largely evidenced study shared with eminent subscribers to different theories, Barrett et al (2019) held two hypotheses to scrutiny: (1) that emotion categories, like fear, anger, disgust, joy, sadness, and surprise, can be expressed by unique facial configurations; and (2) that such categories could be inferred solely from their corresponding facial expressions. Their study pointed to three key insights on how emotional expressions could not be regarded as a valid path for academic, social, commercial and computational pursuits. As stated, the limits imposed concern their reliability, specificity and generalizability. In other words, emotion categories are not all the same; they do not hold a unique facial configuration; and they are impacted by cultural and contextual factors (Barrett et al., 2019).

As posited by Barrett (2017a; 2017b), when one experiences an emotion, the process becomes one where first a core affect, taken as a transient alteration in a neurophysiological state, guides basic behavioural displays, such as hunger or sleep. How one perceives that affect is shaped by a probabilistic computation of value, that is, emotions become a situated, contextualized concept that is in line with the view of cognition as socially situated. Therein lies the reason for heterogeneity in concept formation; individuals construct their conceptual knowledge based on their context and the demands at hand. This means that the situation one is in determines how the emotion enacted is represented; it ties together meaning, context, and experience that eventually becomes highly flexible. Therefore, emotions become situational constructs (Barrett, 2006; 2017a; 2017b).

Taken as situated concepts, emotions can be understood as perceptual symbols that simulate sensorimotor states. Those cannot be reduced into pattern classification as they represent conceptual categories that vary individually thereby not sharing universal or patterned features (Barrett, 2020) as postulated herein in Section 1.1. However, we recognize that it takes many instances for such notions to become acquired knowledge. Therefore, let us pause here and make use of some analogies with our senses to illustrate how this comes to fruition.

Take colour for instance; it is presently acknowledged as the way that the brain interprets and labels light when it refracts from artifacts and is captured by our retina (Kandel et al., 2023). Light, in turn, is measured in wavelengths and each band yields a different colour. Let us turn now to emotions. For emotions to form, stimuli is expressed through the coupling of affect and conceptual knowledge as different categories or patterns to yield different emotions. In a similar way, for colour vision to ensue, light is measured in wavelengths that yield different colours. Taking the analogy a little further to illustrate how situational constraints matter, in the same way as birds can capture ultra violet rays that the human retina cannot transduce due to lacking the specialized photoreceptor, different species may hold a diverse set of receptors – automatic, behavioural or experiential – that confer them with different abilities for emotion detection and treatment thus rendering emotions as dependent on specialized processes within each organism as it interacts with contextual variables.

Let us turn now to taste. For taste to exist, there needs to be gustation (the process), and tastants (the stimuli) transduced into molecules that are captured by specialized receptors. A tastant is a chemical, soluble in water, that activates taste receptor cells to produce a taste sensation via activity in dedicated brain pathways (DiLorenzo et al., 2019). A taste sensation is then conferred by basic tastants whereas flavour, which is multisensory, is how the brain experiences taste in combination with other senses (like smell and vision). An affect, like taste, is generated by basic stimuli (a disturbance) in a neurophysiological state, but it needs to be coupled with conceptual knowledge to yield emotion, as much as taste needs coupling with other senses to give flavour.

A final analogy is the sense of touch, which has several different receptors that form the somatosensory system. There are receptors for pressure and vibrations (mechanoreceptors), for changes in temperature (thermoreceptors), for pain (nociceptors), and for the body in the space (proprioceptors). Likewise, emotions have different categories that are all part of emotional systems that get set into motion by different contextual stimuli, that is, they are culture-specific, and highly individualized (value driven). Those categories are triggered via affective regulation (as in anxious states), bodily regulation (as in different heartbeats), sensations (as in physical pain), and cognition (as in the executive control) so that emotions become rudders (Cantor et al., 2021; Fischer & Bidell, 2006) that impact academic performance; confidence; motivation, persistence, self-control, and curiosity (Immordino-Yang & Damasio, 2007; Meyer & Turner, 2006).

Taken together, these senses – made solely to gain better understanding - illustrate how emotions, taken as organizers of prior knowledge formation and conceptualizations, steer processes that drive automatic actions, beliefs and attributions in conscious and unconscious ways. Therefore, emotions are not different from cognition and perception as brain-related processing mechanisms in that they hold no specific locations and should be rather taken as whole-brain states (Barrett, 2006; 2017a; Barrett & Satpute, 2013). This is an important caveat in the search for biomarkers as it calls for novel approaches in finding biological substrates involved in emotion generation.

Further, studies have shown that feelings do not come from interoceptive information directly and the reason for this rests on attention (Barrett et al., 2004). It is only when internal information is perceived - because it was attended to - that attribution of emotional feelings is possible. However, this is not straightforward; it varies, both among people and even for the same person (Barrett et al., 2004). Thus, biomarkers should not be taken as a silver bullet in the quest for emotions.

*Some Thresholds*

Having briefly addressed the role of biomarkers for emotions and finding that the term is inappropriate as there are no single structures, patterns nor regions that could altogether represent shared instances (Barrett, 2016), let us turn our attention to new research areas, i.e., thresholds. To that end, we need to refer to two crucial actions that are employed when we process an emotion: reverse inference and conditional probability (Barrett et al., 2019). As previously explained in this chapter, the first refers to what we assume based on the best available evidence, e.g., we see a scowl and assume the scowler is angry. The second refers to what we attribute as most likely resulting from a given stimulus e.g., the scowler is most probably angry because of their scowl (Barrett et al., 2019).

Taking emotions as categories of a highly flexible system seems a promising way to understand depressive states, a specially relevant threshold for clinical perspectives. The Bayesian Predictive Coding framework (see Section 1.1) has shown how we are constantly making inferences about incoming input to predict, inform and update, via the generating of models, our knowledge of the world and of ourselves. Doing that is possible via perceptual inference, that feeds into mechanisms of inference as a necessary component (Howhy, 2016), or via active inference, that relies on the body being present, involved in the constant updating of internal models (Clark, 2013). However, if one detaches oneself from such instances, what may unfold is a vicious circle of self-predictions that lock oneself in a confirmation bias akin to a self-fulfilling prophecy generating disengagement and lack of interest (Smith at al., 2020). Also, if one feels overwhelmed and lacking in cognitive appraisal capacity, a negative feedback loop may take hold and lock one in further depressive states (Hutchinson-Wong et al., 2025). What follows is that the depressed individual tends to steer further away from instances of active engagement with stimuli – objects, events, other people and their own perceptions of emotions – thereby closing in on themselves and their appetite for social contact (Giurgi-Oncu et al., 2021).

Another possibility to be explored concerning emotions within the Predictive Processing framework, is how emotions seem to change once one is able to talk and, by talking, reframe thinking about bodily states, such as emotional pain (Solms, 2021). It is by making use of words – an investment in identifying and communicating emotional granularity (Barrett et al., 2004; Tugade, Frederickson & Feldman Barrett, 2004) - that people may find new pathways to comprehend themselves and their states as well as their transience in the face of newer, updated versions of previous processes.

Yet another possibility to explore emotions, in their nature and flexibility, offered by the Bayesian Brain approach, lies with attention in relation to meditation. It is outside the scope of this work to define, characterize and debate the concept of meditation; however, suffice it to say that attention, a cognitive capacity influenced by emotions (Taylor & Fragopanagos, 2005; Yiend, 2010), is impacted by meditation practices (Lippelt, Hommel, & Colzato, 2014; Lutz et al., 2008; Raffone, & Srinivasan, 2010; Sumantry & Stewart, 2021), which relies on a mechanism of balancing acts in the brain (Feldman & Friston, 2010). Attention in a Bayesian framework is taken as the inference about the precision of sensory signals and their causes (Feldman & Friston, 2010). Interestingly, attention in this framework is related to precision and not to the primary cause of sensory input. This leads to a possible exploration of ways to modulate – or balance – the process of synaptic gain that is encoded in precision. As stated in clear terms by Feldman and Friston (2010): “Things get more interesting if we consider that the precision of sensory signals depend on states of the world. This means that optimizing precision entails optimizing inferred states of the world that affect the precision or uncertainty about our sensations.” (op. cit., p. 2). Some stimuli that hold emotional salience can bias attention. This is a bottom-up process - in this case biased - that captures attention and may determine a change in precision (Feldman & Friston, 2010). An exploration of emotional modulation, in turn, may lead to a greater understanding of possible differences in how we experience emotions by relying on mechanisms that encode stimulus based on predictions, errors and balancing acts.

*Closing Remarks*

In this chapter dedicated to emotions, we have proceeded with a tripartite analysis. In the first section we have delineated, in three parts, different directions to explore emotions. First, we introduced defining ideas that have steered notions about what emotions are. Next, we provided a description of different views of emotions across time and showed how research in the area began by considering emotions as innate and homologous. Findings stemming from the early works of Darwin spurred a series of ideas, experiments and models by Hess, Tomkins, Ekman, and Izard. In considering emotions from a physiological perspective, James and Lange presented ideas that found counterpoints in Cannon and Bard and moved the field forward with later models by Schachter and Singer. Ideas put forth by Nauta and Pribram ignited modern notions and models that emphasized the transactional, recursive nature of emotions in body and mind. In a field that deals with dynamical changes of bodily states - or emotions, in short - there is much more to be analysed and considered than facial expressions. Emotions may be common and shared with other species, but that does not turn them into simple concepts or plain, immutable states to be recognized, categorized or mimicked, by artificial intelligence. Care and depth should be the first considerations in how we may explore emotions in different perspectives and for diverse outcomes.

In the second section, we strived to show how recent years has seen a focus on networks in neuroscience – rather than modules – that operationalize how brains perform their task in a metabolically efficient way: they predict bodily states, simulate emotional states, infer actions, and develop heuristics to optimize pathways. In that sense, the constructed theory of emotions stands at the forefront of what evolving research that concerns emotions may reveal. Considering future implications, any field that may come to work with emotions should pay close heed to what this line of research unveils.

In the third section we brought to attention exponential gains in understanding neurobiological substrates of emotion processing through the works of three contemporary researchers, Jaak Panksepp, Antonio R. Damasio and Lisa Feldman Barrett. Rather than pointing towards a single substrate, complex or region that would be akin to a biomarker, their work jointly leads us to regard emotions as whole brain states which, in all their complexity, clearly bind emotions to many substrates, rather than just the one. Underlying the elusive quest for biomarkers, we stressed how understanding the neuronal coding that the brain performs represents a more robust basis upon which thresholds in research on emotional distress, such as in depressive states, and on possible treatments, such as psychoanalysis and meditation, could be found. Stronger efforts to advance integrative methods and models will continue to lead us towards new ways to improve understanding of emotional states and brain function.

*Acknowledgements*

Thanks to Caroline Handley and two anonymous readers for very helpful feedback on earlier drafts of this chapter.

**References**

Aboitiz, F. (2024). *A History of Bodies, Brains, and Minds: The Evolution of Life and Consciousness*. MIT Press.

Adolphs, R., Mlodinow, L., & Barrett, L. F. (2019). What is an emotion?. *Current biology*, *29*(20), R1060-R1064.

Aitchison, L., & Lengyel, M. (2017). With or without you: predictive coding and Bayesian inference in the brain. *Current opinion in neurobiology*, *46*, 219-227.

Alamia, A., & VanRullen, R. (2019). Alpha oscillations and traveling waves: Signatures of predictive coding?. *PLoS Biology*, *17*(10), e3000487.

Amaral, D. G. (2023). Bases Neuroanatômicas pelas quais os Circuitos Neurais Medeiam o Comportamento. In E. R. Kandel et al. (org.). *Princípios de Neurociências*. 6. ed. Porto Alegre: AMGH, 64-83.

Arnold, M.B. (1960), *Emotion and Personality*, New York: Columbia University Press.

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and social psychology review*, *10*(1), 20-46.

Barrett, L. F. (2012). Emotions are real. *Emotion*, *12*(3), 413-429.

Barrett, L. F. (2017a). *How emotions are made: The secret life of the brain.*Pan Macmillan.

Barrett, L. F. (2017b). The theory of constructed emotion: an active inference account of interoception and categorization. *Social cognitive and affective neuroscience*, *12*(1), 1-23.

Barrett, L. F. (2020). *Seven and a Half Lessons about the Brain.* Picador

Barrett, L. F., & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1325-1334.

Barrett, L. F., & Lindquist, K. A. (2008). The embodiment of emotion. IN G.R. Semin and E.R. Smith (Eds.). *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches*, 237-262. New York: Cambridge University Press.

Barrett, L. F., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience: Towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology, 23*(3), 361–372.

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature reviews. Neuroscience*, *16*(7), 419–429. https://doi.org/10.1038/nrn3950

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest, 20*, 1–68. doi:10.1177/1529100619832930

Barrett, L. F., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition & Emotion*, *15*(6), 713-724.

Barrett, L. F., Lewis, M., Haviland-Jones, J. M. (2016). Handbook of Emotions (4th ed.). New York, NY: Guilford Publications, Inc.

Barrett, L. F., Quigley, K. S., Bliss-Moreau, E., & Aronson, K. R. (2004). Interoceptive sensitivity and self-reports of emotional experience. *Journal of personality and social psychology*, *87*(5), 684–697. <https://doi.org/10.1037/0022-3514.87.5.684>

Barrett, L. F. (2016). Navigating the science of emotion. In *Emotion measurement* (pp. 39-84). Woodhead Publishing.

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and economic behavior*, *52*(2), 336-372.

Bengio, Y., Louradour, J., Collobert, R., and Weston, J. (2009). “Curriculum learning,” in *Proceedings of the 26th Annual International Conference on Machine Learning* (Montreal, QC), 41–48.

Berezovskii, V. K., Nassi, J. J., & Born, R. T. (2011). Segregation of feedforward and feedback projections in mouse visual cortex. *The Journal of comparative neurology*, *519*(18), 3672–3683. https://doi.org/10.1002/cne.22675

Berlo, David K. (1960). *The Process of Communication: An Introduction to Theory and Practice*. Holt, Rinehart and Winston. [ISBN](https://en.wikipedia.org/wiki/ISBN_(identifier)) [9780030074905](https://en.wikipedia.org/wiki/Special:BookSources/9780030074905).

Bornstein, M. H., & Lamb, M. E. (2011). Cognitive development. *An Advanced Textbook*. Psychology Press: New York.

Brown, E. C., & Brüne, M. (2012). The role of prediction in social neuroscience. *Frontiers in human neuroscience*, *6*, 147.

Busso, D. S., & Pollack, C. (2015). No brain left behind: Consequences of neuroscience discourse for education. *Learning, Media and Technology*, *40*(2), 168-186.

Buzsáki, G., Logothetis, N., & Singer, W. (2013). Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron*, *80*(3), 751-764.

Calderone, D. J., Lakatos, P., Butler, P. D., & Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends in cognitive sciences*, *18*(6), 300–309. <https://doi.org/10.1016/j.tics.2014.02.005>

Camacho, M. C., Nielsen, A. N., Balser, D., Furtado, E., Steinberger, D. C., Fruchtman, L., Culver, J. P., Sylvester, C. M., & Barch, D. M. (2023). Large-scale encoding of emotion concepts becomes increasingly similar between individuals from childhood to adolescence. *Nature Neuroscience*, *26*(7), 1256–1266. https://doi.org/10.1038/s41593-023-01358-9

Cantor, P., Osher, D., Berg, J., Steyer, L., & Rose, T. (2021). Malleability, plasticity, and individuality: How children learn and develop in context 1. In *The science of learning and development* (pp. 3-54). Routledge.

Castilho, F. M. (2021). A expressão das emoções no homem e nos animais, de Charles Darwin: algumas considerações. *Filosofia e História da Biologia*, *16*(2), 173-207.

Clark, A. (2013). The many faces of precision (Replies to commentaries on “Whatever next? Neural prediction, situated agents, and the future of cognitive science”). *Frontiers in psychology*, *4*, 270

Clark, A. (2016). *Surfing Uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

Clark, A. (2017). Busting out: Predictive brains, embodied minds, and the puzzle of the evidentiary veil. *Noûs*, *51*(4), 727-753.

Clark, A. (2019). Consciousness as generative entanglement. *The Journal of Philosophy*, *116*(12), 645-662.

Clark, A., Friston, K., & Wilkinson, S. (2019). Bayesing qualia: Consciousness as inference, not raw datum. *Journal of Consciousness Studies*, *26*(9-10), 19-33.

Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neuroscience and Biobehavioral Reviews, 112*, 279–299. https://doi.org/10.1016/j .neubiorev.2020.01.032, PubMed: 32018038

Crivelli C., Jarillo S., Russell J. A., Fernandez-Dols J. M. (2016). Reading emotions from faces in two indigenous societies. *Journal of Experimental Psychology: General*, 145(7), 830–843

Crivelli C., Russell J. A., Jarillo S., Fernández-Dols J.-M. (2017a). Recognizing spontaneous facial expressions of emotion in a small-scale society of Papua New Guinea. *Emotion*, 17(2), 337–347

Dalgleish, T. (2004). The emotional brain. *Nature Reviews Neuroscience, 5*, 582-585.

Damasio, A. R. (1994). *Descartes’ Error: Emotion, Reason, and the Human Brain*. Putnam: New York.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1413-1420.

Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. Houghton Mifflin Harcourt.

Damasio, A. (2001). Fundamental feelings. *Nature*, *413*(6858), 781-781.

Damasio, A. (2003). *Looking for Spinoza: Joy, sorrow, and the feeling brain.* Houghton Mifflin Harcourt.

Damasio, A. (2011). Neural basis of emotions. *Scholarpedia, 6*(3), 1804.

Damasio, A. (2021). *Saber e sentir*. São Paulo: Companhia das Letras.

Damasio, A., & Damasio, H. (2023). Feelings are the source of consciousness. Neural *Computation, 35*(3), 277-286.

Damasio, A., Damasio, H., & Tranel, D. (2013). Persistence of feelings and sentience after bilateral damage of the insula. *Cerebral cortex, 23*(4), 833-846.

Damasio, A.R., Tranel, D., Damasio, H. (1991). Somatic markers and the guidance of behavior: Theory and preliminary testing. In: Levin, H.S., Eisenberg, H.M., Benton, A.L. (Eds.), *Frontal Lobe Function and Dysfunction*. Oxford Univ. Press, New York, pp. 217–229

Darwin, C. R. (1872). *The expression of the emotions in man and animals*. London: John Murray.

Davidson, R. J. (2000). Cognitive neuroscience needs affective neuroscience (and vice versa). *Brain and Cognition, 42*(1), 89-92.

Davidson, R. J. (2003). Seven sins in the study of emotion: Correctives from affective neuroscience. *Brain and Cognition, 52*(1), 129-132.

Davidson, R. J., & Sutton, S. K. (1995). Affective neuroscience: The emergence of a discipline. *Current opinion in neurobiology, 5*(2), 217-224.

de Waal FBM, Andrews K. 2022. The question of animal emotions: Do animals, including invertebrates, have felt emotions and does this morally matter? Science 375:1351–52

Dewey, J., (1895). The theory of emotion. II. The significance of emotions. *Psychological Review, 2*, 13–32. doi:10.1037/h0070927

Di Lorenzo, P.M., Chen, JY., Rosen, A.M., Roussin, A.T. (2009). Tastant. In: Binder, M.D., Hirokawa, N., Windhorst, U. (eds) Encyclopedia of Neuroscience. Springer, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-540-29678-2_5888>

Dixon, T. (2003). *From passions to emotions: The creation of a secular psychological category*. Cambridge, UK: Cambridge University Press.

Dixon, T. (2012). “Emotion”: The history of a keyword in crisis. *Emotion Review*, *4*(4), 338-344.

Doelling, K. B., & Assaneo, M. F. (2021). Neural oscillations are a start toward understanding brain activity rather than the end. *PLoS biology*, *19*(5), e3001234.

Dolan, R. (2003). Feeling emotional. *Nature* *421*, 893–894.

Doya, K. (Ed.). (2007). *Bayesian brain: Probabilistic approaches to neural coding*. MIT press.

Dukes, D., Abrams, K., Adolphs, R., Ahmed, M. E., Beatty, A., Berridge, K. C., ... & Sander, D. (2021). *The rise of affectivism. Nature human behaviour, 5*(7), 816-820.

Durán, J. I., & Fernández-Dols, J. M. (2021). Do emotions result in their predicted facial expressions? A meta-analysis of studies on the co-occurrence of expression and emotion. *Emotion*, *21*(7), 1550.

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the national academy of sciences*, *98*(24), 13763-13768.

Ekman, P. (2012). *Emotions revealed: Understanding faces and feelings*. Hachette UK.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in human neuroscience*, *4*, 215.

Fermin, A. S., Friston, K., & Yamawaki, S. (2022). An insula hierarchical network architecture for active interoceptive inference. *Royal Society Open Science, 9*(6), 220226.

Fienberg, S. E. (2006). When Did Bayesian Inference Become "Bayesian"?. *Bayesian Analysis*, *1* (1): 1–40.

Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action and thought. *Handbook of child psychology, 1*, 313-399.

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, *115*(27), E6106-E6115.

Freese, J. L., & Amaral, D. G. (2009). Neuroanatomy of the primate amygdala. In P. J. Whalen & E. A. Phelps (Eds.), *The human amygdala* (pp. 3–42). New York, NY: Guilford Press.

Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, *16*(9), 1325-1352.

Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society of London*. Series B, Biological sciences 360(1456): 815-36.

Friston, K. (2008). Hierarchical models in the brain. *PLoS computational biology*, *4*(11), e1000211.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences* 13(7): 293-301.

Friston, K. (2010). The free-energy principle: a uniﬁed brain theory? *Nature Reviews*. Neuroscience 11(2): 127-138.

Friston, K. (2011). Embodied inference: Or I think therefore I am, if I am what I think. *The implications of embodiment (Cognition and Communication)*, 89-125.

Friston, K. (2012). The history of the future of the Bayesian brain. *NeuroImage*, *62*(2), 1230-1233.

Friston, K. (2014). Active inference and agency. *Cognitive Neuroscience*, *5*(2), 119-121.

Friston, K. J. (2019). Waves of prediction. *PLoS biology*, *17*(10), e3000426.

Friston, K. J., & Frith, C. D. (2015). Active inference, communication and hermeneutics. *Cortex*, *68*, 129-143.

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

Gardner, E. P. & Gardner, D. (2023). Codificação Sensorial. In E. R. Kandel et al. (org.). *Princípios de Neurociências*. 6. ed. Porto Alegre: AMGH, 346-365.

Gendron, M. (2010). Defining emotion: A brief history. *Emotion Review*, *2*(4), 371-372.

Gendron, M., & Barrett, L. F. (2017). Facing the past: A history of the face in psychological research on emotion perception. In J.-M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp.15–36). Oxford University Press.

Gendron, M., & Feldman Barrett, L. (2009). Reconstructing the past: A century of ideas about emotion in psychology. *Emotion review*, *1*(4), 316-339.

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion*, *12*(2), 314.

Gendron M., Crivelli C., Barrett L. F. (2018). Universality reconsidered: Diversity in making meaning of facial expressions. *Current Directions in Psychological Science*, 27(4), 211–219.

Gendron M., Hoemann K., Crittenden A. N., Mangola S. M., Ruark G. A., Barrett L. F. (2020). Emotion perception in Hadza hunter-gatherers. *Scientific Reports*, 10(1), Article 3867.

Giurgi-Oncu, C., Bredicean, C., Frandeș, M., Enătescu, V., Papavă, I., Riviș, I., & Ursoniu, S. (2021). Social Inferences as Mediators of Wellbeing in Depression. *Neuropsychiatric Disease and Treatment*, 1679-1687.

Graeff, F. G. (2018). Evolução da neurociência comportamental nos últimos 50 anos. Foco no sistema de defesa e na neurotransmissão. *Medicina (Ribeirão Preto)*, *51*(2), 89-111.

Hackel, L. M., Larson, G. M., Bowen, J. D., Ehrlich, G. A., Mann, T. C., Middlewood, B., Roberts, I. D., Eyink, J., Fetterolf, J. C., Gonzalez, F., Garrido, C. O., Kim, J., O'Brien, T. C., O'Malley, E. E., Mesquita, B., & Barrett, L. F. (2016). On the neural implausibility of the modular mind: Evidence for distributed construction dissolves boundaries between perception, cognition, and emotion. *The Behavioral and brain sciences*, *39*, e246. <https://doi.org/10.1017/S0140525X15002770>

Headley, D. B., & Paré, D. (2013). In sync: gamma oscillations and emotional memory. *Frontiers in behavioral neuroscience*, *7*, 170.

Hebb, D. O. (1946). Emotion in man and animal: an analysis of the intuitive processes of recognition. *Psychological Review*, *53*(2), 88-106.

Heilbron, M., & Chait, M. (2018). Great expectations: is there evidence for predictive coding in auditory cortex?. *Neuroscience*, *389*, 54-73.

Hess, W. R., & Brügger, M. (1981). The subcortical center of the affective defense reaction. In *Biological order and brain organization: Selected works of WR Hess* (pp. 183-202). Berlin, Heidelberg: Springer Berlin Heidelberg.

Hobson, J. A., & Friston, K. J. (2012). Waking and dreaming consciousness: neurobiological and functional considerations. *Progress in neurobiology*, *98*(1), 82-98.

Hoemann, K., Gendron, M., Crittenden, A. N., Mangola, S. M., Endeko, E. S., Dussault, È., Barrett, L. F., & Mesquita, B. (2023). What We Can Learn About Emotion by Talking With the Hadza. *Perspectives on Psychological Science*, *19*(1), 173-200. <https://doi.org/10.1177/17456916231178555> (Original work published 2024).

Hoemann, K., Wormwood, J. B., Barrett, L. F., & Quigley, K. S. (2023). Multimodal, idiographic ambulatory sensing will transform our understanding of emotion. *Affective Science*, *4*(3), 480-486.

Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Frontiers in Psychology*, *3*, 96.

Hohwy, J. (2016). The self‐evidencing brain. *Noûs*, *50*(2), 259-285.

Hohwy, J. (2020). New directions in predictive processing. *Mind & Language*, *35*(2), 209-223.

Holmes, J., & Nolte, T. (2019). “Surprise” and the Bayesian brain: implications for psychotherapy theory and practice. *Frontiers in Psychology*, *10*, 592.

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological bulletin*, *141*(4), 901.

Huang, Y., & Rao, R. P. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 580-593.

Hutchinson-Wong, N., Glue, P., Adhia, D., & de Ridder, D. (2025). How does depressive cognition develop? A state-dependent network model of predictive processing.*Psychological Review, 132*(2), 442–469. https://doi.org/10.1037/rev0000512

Immordino‐Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, brain, and education*, *1*(1), 3-10.

Immordino-Yang, M. H., Yang, X. F., & Damasio, H. (2016). Cultural modes of expressing emotions influence how emotions are experienced. *Emotion*, *16*(7), 1033.

Izard, C. E. (1977). *Human emotions*. Plenum Press. https://doi.org/10.1007/978-1-4899-2209-0

Izard, C. E. (2010a). The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emotion Review, 2*, 363–370.

Izard, C. E. (2010b). More meanings and more questions for the term “emotion*.” Emotion Review, 2*, 383–385.

Jacobson, H., Ongil, Z., Algom, D., & Usher, M. (2024). Valence in perception: Are affective valence and visual brightness integral dimensions in visual experience?. *Consciousness and Cognition*, *126*, 103783.

James, W. (1884). What is an emotion? *Mind, 9*, 188–205.

Jones, E. G. (2000). Microcolumns in the cerebral cortex. *Proceedings of the National Academy of Sciences*, *97*(10), 5019-5021.

Kalateh, S., Estrada-Jimenez, L. A., Hojjati, S. N., & Barata, J. (2024). A systematic review on multimodal emotion recognition: building blocks, current state, applications, and challenges. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2024.3430850>

Kandel, E.R., Koester J.D., Mack S.H., & Siegelbaum S.A. (2023). *Princípios de Neurociências*, 6e. McGraw Hill.

Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive processing: a canonical cortical computation. *Neuron*, *100*(2), 424-435.

Korzybski, A. (1958). *Science and sanity: An introduction to non-Aristotelian systems and general semantics*. Institute of GS.

Kriegeskorte, N., & Douglas, P. K. (2018). Cognitive computational neuroscience. *Nature neuroscience*, *21*(9), 1148-1160.

Kryklywy, J. H., Ehlers, M. R., Anderson, A. K., & Todd, R. M. (2020). From architecture to evolution: multisensory evidence of decentralized emotion. *Trends in cognitive sciences*, *24*(11), 916-929.

Kuppens, P. (2015). It’s about time: A special section on affect dynamics. *Emotion Review*, *7*(4), 297-300.

Lebrecht, S., Bar, M., Barrett, L. F., & Tarr, M. J. (2012). Micro-valences: Perceiving affective valence in everyday objects. *Frontiers in psychology*, *3*, 107.

LeDoux J. 2021. What emotions might be like in other animals. Curr. Biol. 31(13):R824–29.

Lindquist, K. A., & Gendron, M. (2013). What’s in a word? Language constructs emotion perception. *Emotion Review*, *5*(1), 66-71.

Lippelt, D. P., Hommel, B., & Colzato, L. S. (2014). Focused attention, open monitoring and loving kindness meditation: effects on attention, conflict monitoring, and creativity–A review. *Frontiers in psychology, 5,* 1083.

LoBue, V., and Ogren, M. (2022). How the emotional environment shapes the emotional life of the child. *Policy Insights Behav. Brain Sci.* 9, 137–144. doi: 10.1177/23727322211067264

Lutz, A., Slagter, H. A., Dunne, J. D., & Davidson, R. J. (2008). Attention regulation and monitoring in meditation. *Trends in cognitive sciences, 12*(4), 163-169.

Macrine, S. L., & Fugate, J. M. (2022). Embodied cognition and its educational significance. *Movement matters: How embodied cognition informs teaching and learning*, 13-24.

Malezieux, M., Klein, A. S., & Gogolla, N. (2023). Neural circuits for emotion. *Annual review of neuroscience*, *46*(1), 211-231.

Marblestone, A. H., Wayne, G., & Kording, K. P. (2016). Toward an integration of deep learning and neuroscience. *Frontiers in computational neuroscience*, *10*, 215943.

McEwen, B. S., & Wingfield, J. C. (2010). What is in a name? Integrating homeostasis, allostasis and stress. *Hormones and behavior*, *57*(2), 105–111. https://doi.org/10.1016/j.yhbeh.2009.09.011

McLeod, P., and Dienes, Z. (1996). Do fielders know where to go to catch the ball or only how to get there? *J. Exp. Psychol. Hum. Percept. Perform.* 22, 531.

Mesquita, B., & Walker, R. (2003). Cultural differences in emotions: A context for interpreting emotional experiences. *Behaviour research and therapy*, *41*(7), 777-793.

Mesquita, B., Boiger, M., & De Leersnyder, J. (2017). Doing emotions: The role of culture in everyday emotions. *European Review of Social Psychology*, *28*(1), 95-133.

Meyer, D. K., & Turner, J. C. (2006). Re-conceptualizing emotion and motivation to learn in classroom contexts*. Educational Psychology Review, 18*, 377-390.

Miskovic, V., Kuntzelman, K., Chikazoe, J., & Anderson, A. K. (2016). Representation of affect in sensory cortex. *Behavioral and Brain Sciences*, *39*.

Morris, J. S., Friston, K. J., Büchel, C., Frith, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain: a journal of neurology, 121*(1), 47-57.

Palmer, C. J., Seth, A. K., & Hohwy, J. (2015). The felt presence of other minds: Predictive processing, counterfactual predictions, and mentalising in autism. *Consciousness and Cognition*, *36*, 376-389.

Panksepp, J. (1998). *Affective Neuroscience.* Oxford University Press.

Panksepp, J. (2001). On the subcortical sources of basic human emotions and the primacy of emotional–affective (action–perception) processes in human consciousness. *Evolution and cognition, 7*(2), 134-140.

Panksepp, J. (2003). Damasio's error?. *Consciousness and Emotion, 4*(1), 111-134.

Panksepp, J., & Watt, D. (2011). What is basic about basic emotions? Lasting lessons from affective neuroscience. *Emotion review, 3*(4), 387-396.

Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active inference: the free energy principle in mind, brain, and behavior*. MIT Press.

Peelen, M. V., Berlot, E., & de Lange, F. P. (2024). Predictive processing of scenes and objects. *Nature Reviews Psychology*, *3*(1), 13-26.

Perrinet, L. U. (2020). From the retina to action: Dynamics of predictive processing in the visual system. *The Philosophy and Science of Predictive Processing*, 85-104.

Pessoa, L. (2022). The entangled brain: How perception, cognition, and emotion are woven together. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/14636.001.0001>

Pezzulo, G., Parr, T., & Friston, K. (2022). The evolution of brain architectures for predictive coding and active inference. *Philosophical Transactions of the Royal Society B*, *377*(1844), 20200531.

Picard, F., & Friston, K. (2014). Predictions, perception, and a sense of self. *Neurology*, *83*(12), 1112-1118.

Piekarski, M. (2021). Understanding predictive processing. A review. *AVANT. Pismo Awangardy Filozoficzno-Naukowej*, (1), 1-48.

Platt, M. L., & Spelke, E. S. (2009). What can developmental and comparative cognitive neuroscience tell us about the adult human brain?. *Current Opinion in Neurobiology*, *19*(1), 1-5

Plutchik, R. (1980). Measurement Implications of a Psychoevolutionary Theory of Emotions. In: Blankstein, K.R., Pliner, P., Polivy, J. (eds) *Assessment and Modification of Emotional Behavior. Advances in the Study of Communication and Affect*, vol 6. Springer, Boston, MA. <https://doi.org/10.1007/978-1-4684-3782-9_3>

Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist, 89*(4), 344-350.

Popov, T., Steffen, A., Weisz, N., Miller, G. A., & Rockstroh, B. (2012). Cross‐frequency dynamics of neuromagnetic oscillatory activity: two mechanisms of emotion regulation. *Psychophysiology*, *49*(12), 1545-1557.

Raffone, A., & Srinivasan, N. (2010). The exploration of meditation in the neuroscience of attention and consciousness. *Cognitive Processing, 11*, 1-7.

Ramsay, D. S., & Woods, S. C. (2014). Clarifying the roles of homeostasis and allostasis in physiological regulation. *Psychological review*, *121*(2), 225–247. https://doi.org/10.1037/a0035942

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, *2*(1), 79-87.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology, 39*(6), 1161.

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of personality and social psychology*, *76*(5), 805.

Satpute, A. B., & Lindquist, K. A. (2021). At the neural intersection between language and emotion. *Affective Science*, *2*(2), 207-220.

Saxbe, D. E., Yang, X. F., Borofsky, L. A., & Immordino-Yang, M. H. (2013). The embodiment of emotion: language use during the feeling of social emotions predicts cortical somatosensory activity. *Social cognitive and affective neuroscience*, *8*(7), 806-812.

Scarantino, A. & de Sousa,R. (2021). "Emotion". In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*(Summer 2021 Edition), Available at <https://plato.stanford.edu/archives/sum2021/entries/emotion/> .

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological review*, *69*(5), 379.

Seok, B. (2021).The Emotions in Early Chinese Philosophy. *The Stanford Encyclopedia of Philosophy*(Winter 2021 Edition), Edward N. Zalta (ed.). Availabe at https://plato.stanford.edu/archives/win2021/entries/emotions-chinese/.

Seth, A. K. (2014). A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia. *Cognitive Neuroscience*, *5*(2), 97-118.

Seth, A. K., & Friston, K. J. (2016). Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160007.

Shenhav, A., Barrett, L. F., & Bar, M. (2013). Affective value and associative processing share a cortical substrate. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(1), 46-59.

Siegel, E. H., Wormwood, J. B., Quigley, K. S., & Barrett, L. F. (2018). Seeing what you feel: Affect drives visual perception of structurally neutral faces. *Psychological science*, *29*(4), 496-503.

Singh, A., Westlin, C., Eisenbarth, H., Losin, E. A. R., Andrews-Hanna, J. R., Wager, T. D., ... & Erdogmus, D. (2021, November). Variation is the norm: Brain state dynamics evoked by emotional video clips. In 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 6003-6007). IEEE.

Smith, R., Kuplicki, R., Feinstein, J., Forthman, K. L., Stewart, J. L., Paulus, M. P., ... & Khalsa, S. S. (2020). A Bayesian computational model reveals a failure to adapt interoceptive precision estimates across depression, anxiety, eating, and substance use disorders. *PLoS computational biology, 16*(12), e1008484.

Solms, M. (2021). *The hidden spring: A journey to the source of consciousness*. Profile books.

Spektrum der Wissenschaft (2012). A 1-hour interview with Jaak Panksepp, the father of affective Neuroscience [video]. Available at <https://youtube.com/watch?v=u4ICY6-7hJo%3Fcolor%3Dred%26modestbranding%3D1%26showinfo%3D0%26origin%3Dhttps%3A>

Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain and cognition*, *112*, 92-97.

Sumantry, D., & Stewart, K. E. (2021). Meditation, mindfulness, and attention: A meta-analysis. *Mindfulness, 12*, 1332-1349.

Symons, A. E., El-Deredy, W., Schwartze, M., & Kotz, S. A. (2016). The functional role of neural oscillations in non-verbal emotional communication. *Frontiers in Human Neuroscience*, *10*, 239.

Taylor, J. G., & Fragopanagos, N. F. (2005). The interaction of attention and emotion. *Neural networks, 18*(4), 353-369.

Thut, G., Miniussi, C., & Gross, J. (2012). The functional importance of rhythmic activity in the brain. *Current Biology*, *22*(16), R658-R663.

Todd, R. M., Miskovic, V., Chikazoe, J., & Anderson, A. K. (2020). Emotional objectivity: Neural representations of emotions and their interaction with cognition. *Annual review of psychology*, *71*(1), 25-48.

Tomkins, S. S. (1962). *Affect imagery consciousness: Vol. I. The positive affects.* New York: Springer.

Tomkins, S. S. (1963). *Affect imagery consciousness: Vol. II. The negative affects*. New York: Springer.

Tomkins, S. S. (1991). *Affect imagery consciousness: Vol. III. The negative affects: Anger and fear*. New York: Springer.

Tomkins, S. S. (1992). *Affect imagery consciousness: Vol. IV. Cognition: Duplication and transformation of information*. New York: Springer.

Toronchuk, J. A., & Ellis, G. F. (2013). Affective neuronal selection: The nature of the primordial emotion systems. *Frontiers in Psychology, 3*, 589.

Tscshantz A, Millidge B, Seth AK, Buckley CL (2023) Hybrid predictive coding: Inferring, fast and slow. PLoS Comput Biol 19(8): e1011280. <https://doi.org/10.1371/journal.pcbi.1011280>

Tugade, M. M., Fredrickson, B. L., & Feldman Barrett, L. (2004). Psychological resilience and positive emotional granularity: Examining the benefits of positive emotions on coping and health. *Journal of personality*, *72*(6), 1161-1190.

Van Essen, D. C., Smith, J., Glasser, M. F., Elam, J., Donahue, C. J., Dierker, D. L., ... & Harwell, J. (2017). The brain analysis library of spatial maps and atlases (BALSA) database. *Neuroimage*, *144*, 270-274.

Vignesh, D., He, S., & Banerjee, S. (2024). A review on the complexities of brain activity: insights from nonlinear dynamics in neuroscience. *Nonlinear Dynamics*, 1-22.

Wang, Y., Lu, L., & Wu, M. (2025). Statistical learning across cognitive and affective domains: a multidimensional review. *Frontiers in Integrative Neuroscience*, *19*, 1460471.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of personality and social psychology*, *76*(5), 820.

Westlin, C., Theriault, J. E., Katsumi, Y., Nieto-Castanon, A., Kucyi, A., Ruf, S. F., ... & Barrett, L. F. (2023). Improving the study of brain-behavior relationships by revisiting basic assumptions. *Trends in cognitive sciences*, *27*(3), 246-257.

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and emotion, 24*(1), 3-47.

1. Heuristics are herein referred as priors (prior knowledge) used at large as a way to simplify motor planning and control (using the same example found in McLeod & Dienes, 1996) provided in Marblestone, Wayne, & Kording (2016) [↑](#footnote-ref-1)
2. The Cambridge Dictionary definition of ‘summary’ is of a “a [short](https://dictionary.cambridge.org/dictionary/english/short), [clear](https://dictionary.cambridge.org/dictionary/english/clear) [description](https://dictionary.cambridge.org/dictionary/english/description) that gives the [main](https://dictionary.cambridge.org/dictionary/english/main) [facts](https://dictionary.cambridge.org/dictionary/english/fact) or [ideas](https://dictionary.cambridge.org/dictionary/english/idea) about something”. Available at <https://dictionary.cambridge.org/dictionary/english/summary>. Accessed on August 7, 2025. [↑](#footnote-ref-2)
3. Of note, we take perception in this work as different from behavior according to the process of active inferencing where “perception can be understood as resolving (exteroceptive) prediction errors by selecting predictions that best explain sensations, while behaviour suppresses (proprioceptive) prediction error by changing (proprioceptive) sensations.” (Seth & Friston, 2016, p. 3) [↑](#footnote-ref-3)
4. In 1933, Alfred Korzybski wrote the book "Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics". The work presents his reasoning and logical constructions on the "maps" that encompass all the formulations that humans make. On p. 58 (5th edition), he wrote: "A map is not the territory it represents, but, if correct, it has a *structure similar* to the territory, which explains its usefulness". It is to this reasoning and intended usefulness that we allude to when referring to  *map* and *territory* in the present work.

   [↑](#footnote-ref-4)
5. The Amazing Race is an adventure competition formatted as a reality show. Teams of two people race other teams in tasks where they have to deduce clues, navigate foreign terrain, interact with local people, and accomplish challenges. Many means of transportation lead teams in a progressive elimination route until only a few remain. The first team to get to the finish line is granted the grand prize. Source: <https://en.wikipedia.org/wiki/The_Amazing_Race> [↑](#footnote-ref-5)
6. Barrett (2017b) refers to this mechanism as *allostasis*. For the purposes of general understanding of the present work, *allostasis*, together with its counterpart *homeostasis*, may be better understood as two sides of the same construct: physiological regulation, albeit research in the area does not show total assent (Ramsay & Woods, 2014). While homeostasis refers to how organisms maintain internal balance dynamically; allostasis refers to how an organism's internal parameters vary when change – brought about by internal or external demands – happens (McEwen & Wingfield, 2010). [↑](#footnote-ref-6)
7. Fermin, Friston & Yamawaki (2022) propose a role for the insular cortex in interoceptive self-modeling aligned with the Active Inference framework. They explain how insular interoceptive predictive functions consider the parallel connections of the insular cortex with dorsolateral and ventromedial PFC together with the supplementary motor area (SMA) and the striatum-dopaminergic system for (i) interoceptive habit; or supervised interoceptive control; (ii) interoceptive model-based inference; for categorical and episodic interoception; and (iii) interoception exploration; for Pavlovian and innate interoception responses. Their model highlights the forward connections involving (from lower to higher systems): visceral, subcortical, insula, and PFC (with SMA) to represent interoceptive prediction errors, and the backward connections; aligning the same systems in descending order (from higher to lower), to code for interoceptive predictions. [↑](#footnote-ref-7)
8. According to Barrett (2017b), the study of patterns in emotions can be misguided if not attentive to what patterns mean as in “(...)A pattern that diagnoses sadness is not the brain state for sadness but merely a statistical summary of a highly variable set of instances.” (op. cit., p. 15) [↑](#footnote-ref-8)