The predictive brain: understanding depression through allostasis

CONTENTS

INTRODUCTION

Three common assumptions, although very rooted and often implicit in neuroscience, must be revised in order to better explain both new and past experimental data. A more stable and fitting starting point is crucial to then explain the brain's activity, especially in relation to abnormal states such as depression.

1. Whole-brain signals contribute to mental events: instances of a psychological category originate from activity throughout the brain (Westlin et al., 2023) (the localization assumption would assume otherwise: instances of the same category would be caused by a single, dedicated process implemented in a dedicated neural ensemble).
2. Many neural ensembles for one psychological category: there are (at least) degenerate (many-to-one) mappings between neural ensembles and a psychological category (Westlin et al., 2023) as opposed to the one-to-one assumption which expects that one psychological category is mapped by its dedicated neural pathway[1].
3. An instance of a psychological category emerges from a complex ensemble of signals from the brain, body, and world (Westlin et al., 2023). These signals can only be understood in relation to the rest of the ensemble as opposed to the independence assumption which expects that a function works independently of external, internal, and contextual factors.

The following chapter will explain first how the predictive processing framework emerges as a theoretical consequence of combining both those revised assumptions and a holistic view of data coming from several fields such as computation, evolutionary biology and more. The subsequent section will provide neurobiophysiological experimental data supporting the theory convergently gathered from both humans and

[1]Although the dissertation does not focus on the philosophy of science, it's important to note that this revised assumption is a topic of discussion among some authors. This section only aims to illustrate how the conventional notion is overcome and the more modern one lays the foundation for understanding why the circuits that are discussed later appear to be in charge of numerous distinct processes.

other animals. The aforementioned concepts will then be combined to give a simple but effective look into depression: a mental health disorder that is as prevalent and debilitating as it is challenging to both explain and manage. The insights gained from this research may ultimately inform the development of novel interventions that accurately target the core processes underlying this complex disorder in a more evidence-based manner.

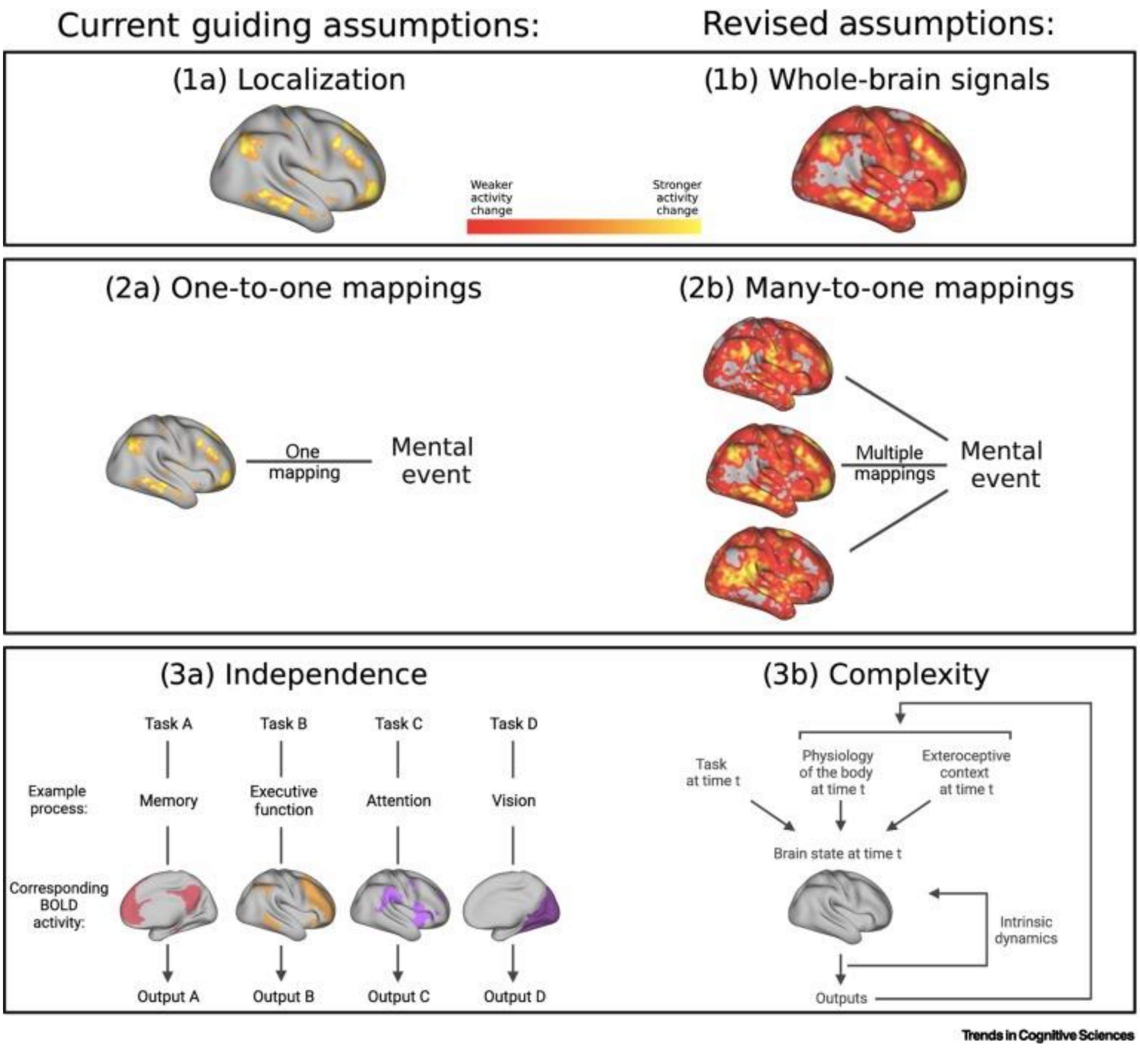

Figure 1. Schematic representation of the traditional and reviewed assumptions (Westlin et al., 2023).

PREDICTIVE PROCESSING FRAMEWORK

Core elements and hypotheses

In its simplest terms the predictive processing approach uses a form of Bayesian inference to navigate the world (Barrett et al., 2016; Hutchinson & Barrett, 2019; Kube et al., 2020; Seth & Friston, 2016): the brain is continually running an internal model of an animal's world which is refined based on comparisons with incoming sensory information from the body and the outside; its intrinsic activity can be either confirmed or modified by comparison with sensory input. The model is generative, meaning that past experiences can be recombined (as predictions, or top-down processing) in novel ways as they are remembered (Hutchinson & Barrett, 2019; Kleckner et al., 2017; Shaffer et al., 2023).

Several studies, discussed in Hutchinson & Barrett, 2019, prove experimentally the temporal dependence: people are experientially blind when presented with what appears to be a random assortment of black and white blobs (*Mooney images*) that in reality are nothing but visually degraded versions of regular images. Following exposure to the natural source images that were used to create them, these images are subsequently interpreted as cohesive. This suggests that exposure to the source images shapes perceptual experience, but not the other way around. Another example shows that when participants are asked to adjust the color of a well-known object such as a banana to be grey, they tend to over-adjust (i.e., they adjust it to be more blueish-grey, which combined with the predicted yellow produces a subjective judgement of gray).

The discrepancies between the expected predictions and the sensory inputs are called prediction errors, in other words: learning signals (or bottom-up processing). Evaluating the importance of those errors (a phenomenon commonly called attention or salience (Katsumi et al., 2022)) to update the internal model, is metabolically expensive: choosing which prediction errors to weigh and how, and then implementing those to update the model, is the core function of a brain and it takes roughly 20% of the total energy consumed (Barrett et al., 2016; Shaffer et al., 2022).

One specific instantiation of this framework is allostasis: the process by which the brain anticipates the needs of the body and attempts to meet those needs before they arise (Barrett et al., 2016; Hutchinson & Barrett, 2019; Katsumi et al., 2022; Kleckner et al., 2017; Shaffer et al., 2022, 2023). Unlike homeostasis, which aims to maintain fixed set points, allostasis involves proactive adjustments to minimize the impact of anticipated stressors and optimize the body's resources. Allostasis is constantly operating, regardless of whether an animal is awake (active) or at rest; it is not a state, but rather a continuous process.

A unified science of the brain, body, and mind may now be built on the discovery that allostasis is a crucial component of the state space of the brain, which functions as a complex, nonlinear, dynamical system that constantly interacts with its surroundings and its own body. Two hypotheses are inferable for psychological science: (1) single mental events do not arise in a vacuum but are temporally dependent on prior events, (2) energy regulation, plus its affective consequences, are core features of all psychological phenomena, not just those that are emotional or involve fight or flight (Barrett et al., 2016; Hutchinson & Barrett, 2019; Kleckner et al., 2017).

Interoception, defined as any bodily information that is sent either via (1) small diameter (unmyelinated) C-fibers or (myelinated) Aδ-fibers, lamina I, the spinothalamic tract and then onto the insula and anterior cingulate cortex (ACC), or (2) cranial nerves (glossopharyngeal and vagus) to the nucleus of the solitary tract (Katsumi et al., 2022; Murphy et al., 2017; Seth & Friston, 2016), is another key actor working for a functioning allostasis. Some of the signals it relays are related to hunger, temperature, heart rate, blood sugar and much more. Notably it doesn't only refer to conscious perceptions of one's states, but also implicit and subconscious (Katsumi et al., 2022). A significant role for interoception in higher-order cognition is supported by empirical evidence demonstrating that interoceptive ability, the accurate perception of internal states, predicts competence in a variety of emotional domains as well as in learning and decision-making and some traditionally physical ones. Within the affective domain, interoception appears to be necessary for all aspects of emotional processing (Murphy et al., 2017; Seth & Friston, 2016). Importantly, interoception has been argued to underpin selfhood and self-awareness and sociocognitive and socioaffective ability (Murphy et al., 2017). Both interoceptive signals and cognitive evaluation of one's internal and external environment contribute to emotional experience. Interoceptive signals do not cause emotional awareness or vice versa, there is a circular causality: neuronally encoded predictions about bodily states engage autonomic reflex through active inference while interoceptive signals inform and update these predictions (Seth & Friston, 2016).

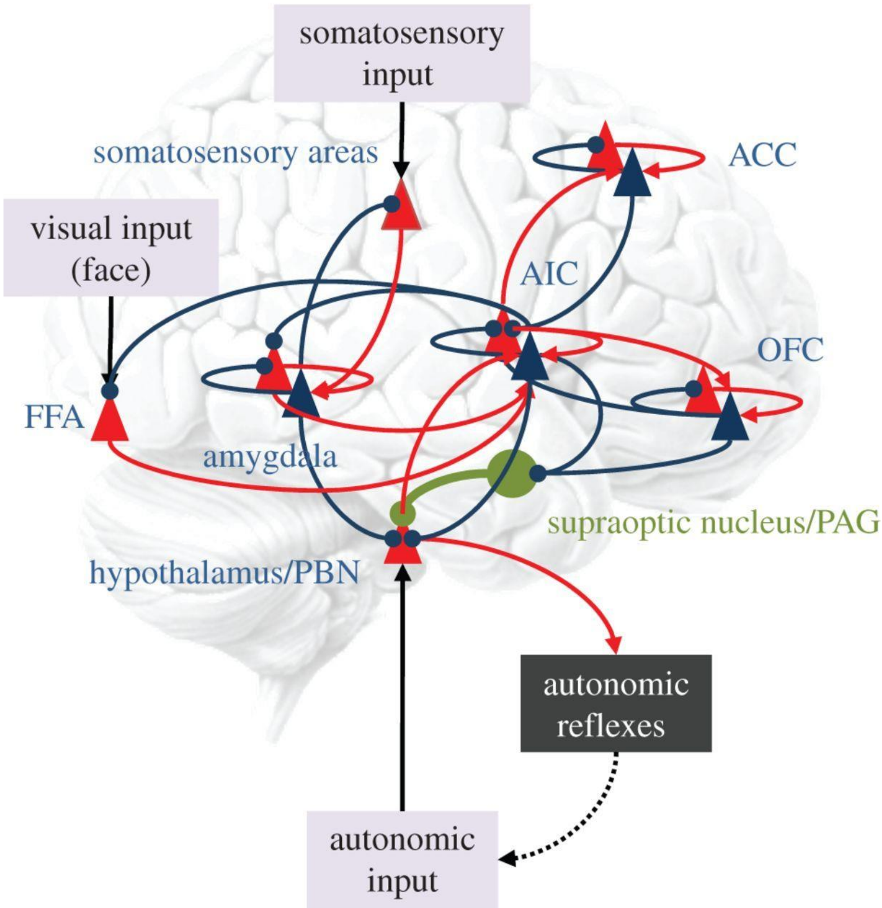

Figure 2. A simplified representation of the encoding of three possible inputs and their predictions. Red triangles represent neurons encoding prediction error, while blue ones represent neurons encoding expectations. Arrows denote excitatory connections, circles inhibitory (Seth & Friston, 2016).

Importantly, interoceptive predictions constitute just one stream of multimodal predictions that are generated by expectations about the embodied self. Furthermore, there is a crucial role to signals of punishment and reward, making the accurate perception and recognition of these signals fundamental to learning, this is one of the many possible processes that become dysfunctional in depression. (Seth & Friston, 2016).

The vagus nerve is also another fundamental actor in the dialogue between brain and body, engaged not only in signal relay but also in signal processing (Shaffer et al., 2023).

Intuitively, such a whole-brain, domain-general task like allostasis, should be supported by at least one equally broad network in the brain. Those networks do in fact exist: the default mode network (DMN) and salience network (SN) work together to create a highly connected functional ensemble for integrating information across the brain, with interoceptive and allostatic information at its core, even though it may not be apparent much of the time (Kleckner et al., 2017). In essence it appears that the DMN runs an internal model of the world, supporting mentalizing and meaning making, together with the hippocampus, generatively constructing the most abstract, compressed, features of the brain's internal model, which are then decompressed into prediction signals. The SN predicts the allostatic relevance of prediction errors, modulating which errors to learn and which to ignore as noise: as the name suggests, salience (Barrett et al., 2016). Evidence for both networks and their interconnection will be provided in the following chapter.

Neurobiophysiological correlates

Decades of studies in macaques show that prediction signals flow from regions with less laminar development (for example agranular regions) to regions with greater laminar development (for example granular regions), whereas prediction error signals flow in the other direction (Kleckner et al., 2017; Murphy et al., 2017; Seth & Friston, 2016; Shaffer et al., 2023).

The bodily signals related to interoception are thought to be represented mainly within the insula (more specifically the anterior insular cortex (AIC)) and anterior cingulate cortex (ACC) but also possibly the subgenual cortex (SGC) and orbitofrontal cortex (OFC) leading these structures to be collectively referred to as the "interoceptive cortex" or visceromotor areas (VMAs) (Kleckner et al., 2017; Murphy et al., 2017; Seth & Friston, 2016). It is noteworthy though, that although they are thought to be where the signals converge, there is a case study of a single patient who, despite their insula damage, had largely intact perceptions of pain and response to tickling, so further analysis may be required (Murphy et al., 2017).

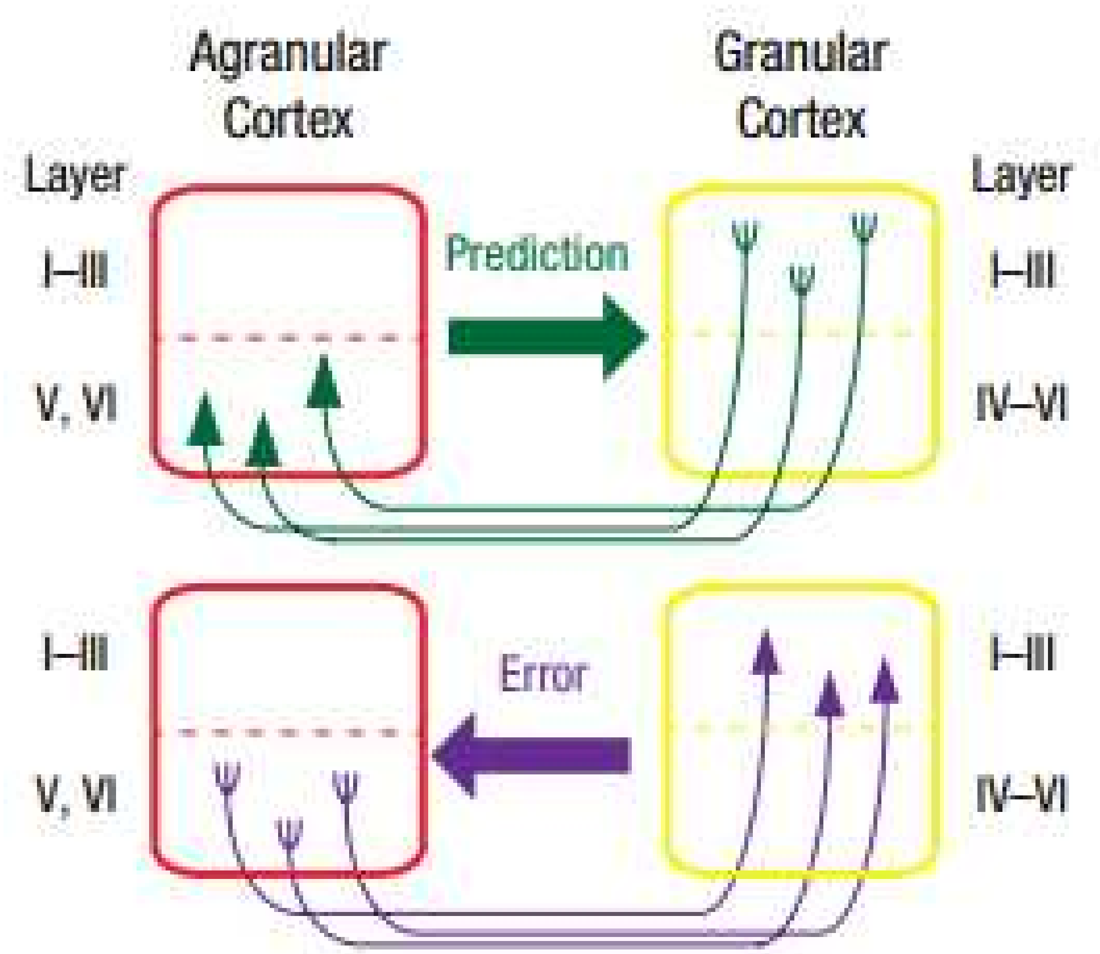

Figure 3. Information flow between cortical regions from Hutchinson & Barrett, 2019.

These areas collectively embody a generative model of interoceptive responses and issue predictions that, when unpacked at the lower hierarchical level, serve as homeostatic set-points. These VMAs are known to receive ascending projections from viscerosensory areas (e.g., posterior and mid-insula) and their descending connections engage a range of subcortical, brainstem and spinal cord targets involved in visceromotor control, such as the periaqueductal grey (PAG) and the parabrachial nucleus (PBN) but also the central nucleus of the amygdala and the central pattern generators of the hypothalamus (Seth & Friston, 2016). As well as known anatomical connectivity patterns, this basic architecture is supported by cytoarchitectonic observations that VMAs lack a well-formed (granular) layer IV and relatively undifferentiated superficial layer (corresponding to II and III, see Figure 3): predictions flow from these cortical structures to ones that are known to control the body's internal milieu (Barrett et al., 2016; Katsumi et al., 2022; Kleckner et al., 2017; Shaffer et al., 2022, 2023). Descending predictions provide a homeostatic set-point against which primary interoceptive afferents can be compared. The resulting prediction error then drives sympathetic or parasympathetic effector systems to ensure allostasis, for example, sympathetic smooth-muscle vasodilatation as a reflexive response to the predicted interoceptive consequences of "blushing with embarrassment" (Seth & Friston, 2016). Simultaneously, the ACC and AIC send the anticipated sensory consequences of those visceromotor actions to the more granular primary interoceptive cortex in the dorsal to mid-posterior insula. As a further confirmation, most of the key visceromotor regions in the system do, in fact, have monosynaptic, bidirectional connections to the primary interoceptive cortex reinforcing the idea that they directly exchange interoceptive prediction signals with each other (Kleckner et al., 2017). It's also confirmation that the DNAs do indeed monosynaptically project to the subcortical and brainstem regions that control the internal milieu (Kleckner et al., 2017). Dopamine is hypothesized to support vigorous action and learning that is necessary to secure the rewards that maintain efficient allostasis rather than playing a necessary or sufficient role in rewards themselves (Barrett et al., 2016). The dopaminergic system is indeed one target of depression treatments.

Instead of thinking of the interoceptive system and the allostatic system as different, they should be thought as part of the same network (Kleckner et al., 2017; Shaffer et al., 2022). This system supports a wide range of psychological phenomena, all explainable by their reliance on allostasis. Regions controlling inner body physiology lie in networks that also support social affiliation, pain, judgements, empathy, reward, addiction, memory, stress, craving among others (Kleckner et al., 2017). More and more studies have found that the SN and DMN are domain-general. Some models consider regions that are part of the DMN to be higher in the processing hierarchy than regions of the SN in the brain. The former is hypothesized to infer the meaning of interoceptive signals based on past experience or issue predictions to VMAs within the SN based on the brain's beliefs about its capacity to successfully perform allostasis (Barrett et al., 2016; Katsumi et al., 2022; Shaffer et al., 2023). These two networks contain the highest proportion of hubs belonging to the brain's "rich club", defined as the most densely interconnected regions in the cortex (see Figure 4). All other sensory and motor networks communicate with the default mode and salience networks, and potentially with one another, through these hubs. Rich-club hubs that are limbic in structure, as opposed to non-limbic, exhibit topological properties more suited to function as "high-level" connectors, integrating already highly integrated information across modules or communities of regions and synchronizing information flow (Barrett et al., 2016; Katsumi et al., 2022; Kleckner et al., 2017). The agranular hubs within the two networks, which are also visceromotor control regions, are the most powerful predictors in the brain. Indeed, hub regions in these networks display a pattern of connectivity that positions them to easily send prediction signals to every other sensory system in the brain. Some studies have found that activities in the DMN and SN have an inverse or negative relationship, or anti-correlated, meaning that as one network increases its neural activity relative to baseline, the other decreases. Such findings have more recently been challenged on both statistical and theoretical grounds. In fact, when global signal is removed in pre-processing, the two networks can show a pattern of positive connectivity (Kleckner et al., 2017). Further confirmation comes from the fact that dorsal mid to posterior insula, which, as said, functions as primary interoceptive cortex, is a point of overlap for these networks, suggesting that processing of unanticipated interoceptive and exteroceptive signals may be influenced by their predictive allostatic relevance (Katsumi et al., 2022).

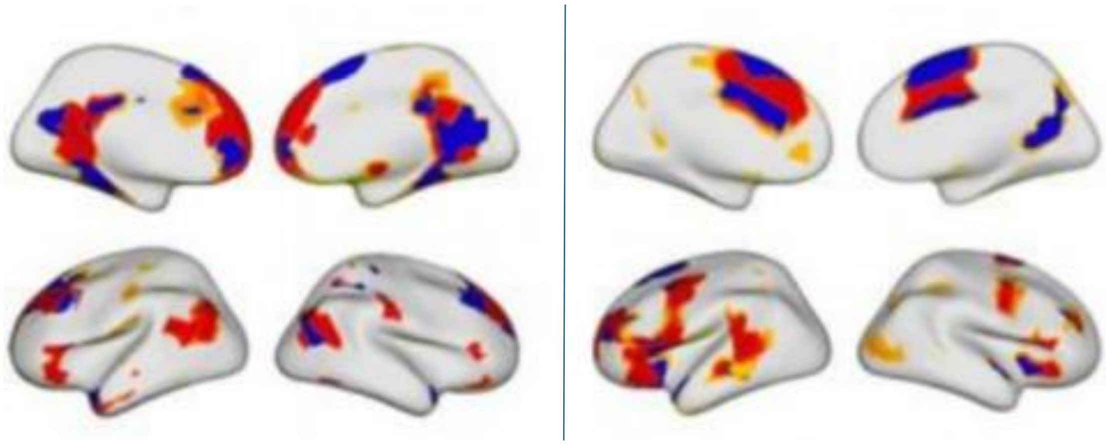

Figure 4. Spatial overlap of rich club (blue) with DMN (left) and SN (right) (van den Heuvel & Sporns, 2013).

Although until now it has been explained how information is generated, received, and elaborated centrally mainly in the DMN and SN, it is also important to focus on how most of the viscerosensory signals are transmitted from the body to said networks. The vagus nerve, consisting of both afferent (sensory, circa 80% of them) and efferent (motor, the other 20%) fibers, plays a key role in the transmission of signals. These fibers seem to share the same fascicles, suggesting opportunities for signals to be influenced by each other before they reach the nodose ganglion or other brainstem targets, a process known as ephaptic coupling (Shaffer et al., 2023). Evidence suggests that the vagus nerve also has the ability to compress signals, which involves removing redundancies to increase signal quality while decreasing processing cost, with a similar gradient as discussed earlier for the cerebral cortex but also the hippocampus and cerebellum. Somatosensory signals are compressed along a gradient as they arise in the dorsal column of the spinal cord, and signal compression may occur in the peripheral nervous system as well (Shaffer et al., 2023). An example of this phenomenon can be found in the enteric nervous system which shows that signal integration happens in the peripheral nervous system too. New methodological approaches also suggest that the afferent vagus is capable of both spatial and temporal co-occurrence mapping, without precise viscerotopy, exactly like the cerebral cortex (Shaffer et al., 2023).

When considered collectively, all these findings point to the fact that allostasis is the central, anatomical, and physiological focus of the brain and body. Different psychological categories may now be re-interpreted as distinct applications and nuances of the same task, instead of separate states arising from unique computations that are localized to specific regions.

DEPRESSION AS METABOLIC DYSREGULATION

Depressive disorder

Depression[2] is a debilitating syndrome that presents with abnormalities in neurologic, metabolic, and immunologic systems, as well as aberrant hypothalamic-pituitary-adrenal (HPA) axis function and pervasive negative affect. The World Health Organization estimates that 5% of adults, or approximately 280 million people in the world, suffer from it (WHO, 2022). Although it is complicated to attribute the exact cost for society due to a large number of factors, such as imprecise direct, indirect, and intangible costs, the economic burden of this disease comes close to $1 trillion yearly (The Lancet Global Health, 2020). Depression also aggravates the course of cardiovascular conditions, it is linked to obesity, osteoporosis, arthritis, type 2 diabetes, certain cancers, periodontal disease, and frailty (Demiralp et al., 2012). Being as prevalent and relevant as it is, it might be expected that a big focus is driven toward finding the most effective treatments possible, but in reality, that is not necessarily the case. A systematic review (Khan et al., 2012) analyzing both approved, not approved, combined, and control treatments found that the combination of psychotherapy and antidepressants may provide a slight advantage to controls or alternative treatments. From that analysis, it also appears that those elements alone, although differing from placebo trials, did not differ from alternative therapies such as exercise and acupuncture. Nevertheless, it’s crucial to underline that individuals participating in these trials do not necessarily represent accurately depressed patients in clinical practice, hence, especially for alternative therapies, the group of depressed individuals is highly specific in traits.

Just as an accurate interoception improves accuracy in higher cognition, in particular emotional experience, there is a differentiated perception of such state between normal people and individuals with depression (Murphy et al., 2017). Among the symptoms of

[2] Although there are some differences that are beyond the scope of this dissertation, throughout the thesis the term depression will be used interchangeably with major depressive disorder, and generally to indicate an abnormal and pathological depressed mood.

depression, there can be found a class of deficits for negative information. Generally speaking, mood disorders fundamentally alter the way in which the self is situated in the physical and social world. People with Major Depressive Disorder (MDD) frequently anticipate negative events or experiences, whereas they rarely anticipate positive events or experiences (Kube et al., 2020). Also, people who lack the ability to differentiate specific negative emotional states from each other or from a general feeling of unpleasantness might choose actions that are not appropriate to the current context and might exacerbate the problem (Demiralp et al., 2012). As a matter of fact, people with depression have less differentiated negative emotions than healthy individuals, while there is no significant difference in differentiation of positive emotions. Notably, a study evaluating exactly these mechanisms, also found no correlation for people in either group (depressed patients or healthy ones) for negative and positive differentiation, suggesting that these two tasks are separate and dependent on different psychological mechanisms (Demiralp et al., 2012). Gender too was not a significant moderator between the variables. Therefore, individuals with major depressive disorder experience negative emotions with less differentiation in their lives than healthy individuals. For the internal model and allostasis, poor differentiation means an inaccurate evaluation of data that then causes an inaccurate response to the stimuli, fueling the depressive cycle. Depression is also associated both with alexithymia, a condition defined by an impoverished conceptual understanding of emotion, and intense negative affect. Importantly, both alexithymia and depression are linked to diminished interoceptive awareness (Barrett et al., 2016). The first key takeaway for psychotherapy is, therefore, focusing on improving interoception accuracy.

In a healthy brain, prediction error signals allow for learning to better tailor the brain's internal model to the immediate circumstances and, thereby, to minimize future error and improve efficient allostasis. If the depressed brain is viewed as "locked in", coupled with persistent negative affect and difficulty in engaging in vigorous mental or physical activity, a relative insensitivity to prediction errors and/or poorly calibrated precision estimates develops. Both problems would lead to a failure of model updating, which, in turn, prompt further inefficiency, producing a downward spiral (rumination is indeed another symptom of depression, and it is also associated with increased connectivity between the sgACC and the rest of the DMN) (Barrett et al., 2016). In other words, in terms of a self-fulfilling prophecy, people with depression perceive their world as being mostly negative because they expect it to be mostly negative. Several lines of research converge on the finding that depressed people have difficulty updating negative expectations after unexpected positive experiences, they tend to interpret ambiguous situations often negatively and less often positively, especially if they contain self-referential stimuli. Moreover, they maintain established negative interpretations of ambiguous information even after novel information that disconfirms the initial negative interpretation in favor of a more positive interpretation (Kube et al., 2020; Ramos-Grille et al., 2022). Additionally, another study (Ramos-Grille et al., 2022), showed empirically that individuals with depression showed lower predictability ratings, i.e., an increased difference between expected facial expressions and displayed facial expression for negative (both sadness and fear) but not positive (happiness) evoked emotions.

Another possible path of development for depression is a generally metabolically inefficient internal model, for example, if the metabolic demand on the body at some point was too large, the brain may not readjust correctly, as is the case for severe adverse childhood experiences such as traumatic events and neglect. It's also the case for persistent presence of low-grade stressors, and even small but numerous challenges or physical illnesses could lead an individual to construct a profoundly negative internal model. The brain samples past experiences to create predictions of the immediate future, so if all the data available to it is generally dangerous, the current context will be affected by previous experiences and subsequent predictions (Barrett et al., 2016; Jungilligens et al., 2022). There is clinical confirmation (Barrett et al., 2016; Demiralp et al., 2012) of such claim: all aforementioned stressors can cause depression, which includes all kinds of metabolic symptoms like unpleasant mood, irritability, anhedonia. At this point, it's clear that symptoms of depression are all expression of inefficient energy regulation, where the misalignment between the brains' anticipation of bodily needs leads to maladaptive psychological responses.

There is clear degeneracy in reaching a depressed state since there are multiple pathways that shift the brain into dysregulated allostasis. An example of many-to-one mappings in depression are, for example, the findings of both increased and decreased resting metabolism in the subgenual portion of the ACC (Barrett et al., 2016). As mentioned earlier, this area is involved in maintaining sympathetic and parasympathetic autonomic control of the viscera, so it becomes evident that either subgenual ACC hyper-activation (fatigue, lethargy) or hypo-activation (agitation, irritation) can promote inefficient allostasis, leading to some of the symptoms of depression. Many other areas, mentioned in the previous chapters as fundamental in building "emotional" activation, happen to indeed be abnormal in depressed patients: amygdala responsivity for example, seems to be increased to negative stimuli in depressed patients, which explains the increased relevance of negative information in such patients. The medial-prefrontal cortex seems to be responsible for rumination and self-referential processing, and indeed depressed patients exhibit increased activity when ruminating compared to healthy controls, in addition to decreased top-down inhibition of these processes. PET studies also show increased resting state sgACC and thalamic functional connectivity with the DMN in individuals with MDD relative to healthy controls (Park et al., 2019). Moreover, depressed individuals also exhibit decreased functional connectivity of subcortical limbic structures, such as the amygdala, to prefrontal cortical regions, while they present with increase in connectivity of the amygdala to the temporal poles. Increased amygdala activity is also usually interpreted to give rise to intense negative mood and rumination (Shaffer et al., 2022). Finally, the dmPFC, described as an integral structure in regulation of communication between the DMN and SN is under activated in depressed patients (Park et al., 2019).

It's relevant to note that there are many forms of depression which are episodic, that might be because there are many ways to disrupt allostasis temporarily: changes in eating, sleeping, exercise or even a loss of a loved one, can lead to transient changes in energy regulation. This is consistent with both the degenerate emergence of depression and with common clinical situations that can lead to the development of such condition.

Available therapies and their focus

The previously presented fact of the degenerate nature of depression is one of the reasons and the marked variance in their findings (Shaffer et al., 2022) is why the condition is so difficult to treat. It's unclear just by analyzing a patient's symptoms, which is the pathway affected, since there is abundant redundancy. At the same time, by suggesting that depression arises from a chronic energy inefficiency and generally altered interoceptive signaling through a computational architecture, several targets of opportunity for intervention and treatment can be recognized. Firstly, depression may be relieved by directly affecting the descending allostatic predictions that originate in agranular limbic cortices: that is the target for deep brain stimulation of the sgACC that achieves antidepressant effects. Secondly, interventions that address the HPA axis dysregulation, such as cortisol synthesis inhibitors and anti-inflammatory medications have shown great promise (Barrett et al., 2016). Thirdly, vagal nerve stimulation may have its efficacy explained in reducing the noisy afferent interoception prediction errors gives the important mediative role of the vagus nerve. Lastly, treatments that offer the opportunity for recategorization provide possibly the most effective intervention pathway: cognitive behavioral therapy (CBT) may have its effects, over time, by helping a person construct new concepts, that, as prediction signals, modify the gain on prediction errors via the salience network. Convergent proof shows that CBT alters the activity of two key regions in the brain's rich club: the cingulate and AICs (Barrett et al., 2016). Another meta-analysis on CBT, comparing its efficacy to other therapies and control conditions among 409 trials and 52,702 patients (Cuijpers et al., 2023), found CBT to have significant and lasting (6-12 months) effects compared to control conditions and although not significantly different from pharmacotherapies at the short term, they were so at 6-12 months follow-up. Interestingly, this paper found no significant effectiveness for combined therapy as opposed to just CBT at either time point. The conclusion states that CBT appears to be effective for depression, even though there is no evident superiority over other forms of psychotherapy.

As mentioned earlier, dopamine is suspected to have a fundamental role in learning and updating predictions, and since cognitive immunization is a key symptom of depression (the difficulty of patients to responding to prediction errors) it's coherent that available treatments also target dopaminergic systems. Finally, even though SSRIs seem to be the first-line antidepressants they are not always effective, and importantly it has been repeatedly shown that serotonin seems to play an important role in metabolism and energy regulation (Shaffer et al., 2022) which further corroborated the energy dysregulation hypotheses (another observation is the fact that side effects of SSRI do show metabolic symptoms such as weight changes and dyslipidemia). Also, they seem to normalize amygdala responsivity to negative stimuli (Park et al., 2019; Ramos-Grille et al., 2022).

Finally, a recent review aiming to find the mediators through which psychotherapy is supposed to work (Lemmens et al., 2016), although expressing important difficulty in finding said mediators (especially in showing causality), noted that indeed several processes such as dysfunctional attitudes, negative thoughts, rumination, worry, and mindfulness skills were associated with change in the majority of studies reviewed.

CONCLUSIONS

It's not surprising that when taking into consideration reviews and meta-analysis, there is no definitive answer to which is the best line of intervention while there are widespread differences. In future research, a holistic approach that combines neurobiological, psychological, and computation methods is crucial. This could involve, among other ideas, longitudinal studies tracking how interventions like CBT affect the brain's predictive models and allostasis in real-time, using neuroimaging techniques and physiological monitoring (some studies suggest, for example, that patients treated with CBT exhibit reduced DMN activity in response to negative stimuli (Park et al., 2019)). It also may include improvements to CBT, updating the cognitive model of depression with more specific interpretations such as work on both the dysfunctional predictions and expectations per se and lack of updating them, as also suggested in Jungilligens et al., 2022.

As shown throughout this thesis, although depression is extremely relevant in society today and will continue to be in the near future, the standard of treatment is not necessarily on par with its gravity. Through the predictive processing framework, depression can be re-characterized behaviorally as a metabolic disease, which is a disruption of allostasis not only on a theoretical level but also empirically testable. The goal is not to reduce every mental phenomenon to energy regulation but rather to highlight energy regulation as a key element of the state of a brain. Understanding the mechanisms of this underlying state of the brain will be crucial to develop more effective treatments, focused on acting on the precise abnormal processes, both through biological and behavioral means. Available and already effective treatments should be further improved by updating them with new knowledge and eliminating unverified claims. Outdated treatments should be revised and overcome in favor of evidence-based once that utilize the most fitting explanations in neuroscience instead of mostly dogmatic approaches. New treatments should start being developed by focusing on the physiology of the body and brain and be thoroughly confirmed by neuroimaging and other techniques.

REFERENCES

Barrett, L. F., Quigley, K. S., & Hamilton, P. (2016). An active inference theory of allostasis and interoception in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160011. https://doi.org/10.1098/rstb.2016.0011

Cuijpers, P., Miguel, C., Harrer, M., Plessen, C. Y., Ciharova, M., Ebert, D., & Karyotaki, E. (2023). Cognitive behavior therapy vs. control conditions, other psychotherapies, pharmacotherapies and combined treatment for depression: a comprehensive meta-analysis including 409 trials with 52,702 patients. *World Psychiatry*, *22*(1), 105–115. https://doi.org/10.1002/wps.21069

Demiralp, E., Thompson, R. J., Mata, J., Jaeggi, S. M., Buschkuehl, M., Barrett, L. F., Ellsworth, P. C., Demiralp, M., Hernandez-Garcia, L., Deldin, P. J., Gotlib, I. H., & Jonides, J. (2012). Feeling Blue or Turquoise? Emotional Differentiation in Major Depressive Disorder. *Psychological Science*, *23*(11), 1410–1416. https://doi.org/10.1177/0956797612444903

Hutchinson, J. B., & Barrett, L. F. (2019). The Power of Predictions: An Emerging Paradigm for Psychological Research. *Current Directions in Psychological Science*, *28*(3), 280–291. https://doi.org/10.1177/0963721419831992

Jungilligens, J., Paredes-Echeverri, S., Popkirov, S., Barrett, L. F., & Perez, D. L. (2022). A new science of emotion: implications for functional neurological disorder. *Brain*, *145*(8), 2648–2663. https://doi.org/10.1093/brain/awac204

Katsumi, Y., Theriault, J. E., Quigley, K. S., & Barrett, L. F. (2022). Allostasis as a core feature of hierarchical gradients in the human brain. *Network Neuroscience*, *6*(4), 1010–1031. https://doi.org/10.1162/netn_a_00240

Khan, A., Faucett, J., Lichtenberg, P., Kirsch, I., & Brown, W. A. (2012). A Systematic Review of Comparative Efficacy of Treatments and Controls for Depression. *PLoS ONE*, *7*(7), e41778. https://doi.org/10.1371/journal.pone.0041778

Kleckner, I. R., Zhang, J., Touroutoglou, A., Chanes, L., Xia, C., Simmons, W. K., Quigley, K. S., Dickerson, B. C., & Feldman Barrett, L. (2017). Evidence for a large-scale brain system supporting allostasis and interoception in humans. Nature Human Behaviour, 1(5), 0069. https://doi.org/10.1038/s41562-017-0069

Kube, T., Schwarting, R., Rozenkrantz, L., Glombiewski, J. A., & Rief, W. (2020). Distorted Cognitive Processes in Major Depression: A Predictive Processing Perspective. *Biological Psychiatry*, *87*(5), 388–398. https://doi.org/10.1016/j.biopsych.2019.07.017

Lemmens, L. H. J. M., Müller, V. N. L. S., Arntz, A., & Huibers, M. J. H. (2016). Mechanisms of change in psychotherapy for depression: An empirical update and evaluation of research aimed at identifying psychological mediators. *Clinical Psychology Review*, *50*, 95–107. https://doi.org/10.1016/j.cpr.2016.09.004

Murphy, J., Brewer, R., Catmur, C., & Bird, G. (2017). Interoception and psychopathology: A developmental neuroscience perspective. *Developmental Cognitive Neuroscience*, *23*, 45–56. https://doi.org/10.1016/j.dcn.2016.12.006

Nord, C. (2024). *The Balanced Brain*. Princeton University Press. https://doi.org/10.1515/9780691259055

Park, C., Rosenblat, J. D., Lee, Y., Pan, Z., Cao, B., Iacobucci, M., & McIntyre, R. S. (2019). The neural systems of emotion regulation and abnormalities in major depressive disorder. *Behavioural Brain Research*, *367*, 181–188. https://doi.org/10.1016/j.bbr.2019.04.002

Ramos-Grille, I., Weyant, J., Wormwood, J. B., Robles, M., Vallès, V., Camprodon, J. A., & Chanes, L. (2022). Predictive processing in depression: Increased prediction error following negative valence contexts and influence of recent mood-congruent yet irrelevant experiences. *Journal of Affective Disorders*, *311*, 8–16. https://doi.org/10.1016/j.jad.2022.05.030

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. https://doi.org/10.1098/rstb.2016.0007

Shaffer, C., Barrett, L. F., & Quigley, K. S. (2023). Signal processing in the vagus nerve: Hypotheses based on new genetic and anatomical evidence. *Biological Psychology*, *182*, 108626. https://doi.org/10.1016/j.biopsycho.2023.108626

Shaffer, C., Westlin, C., Quigley, K. S., Whitfield-Gabrieli, S., & Barrett, L. F. (2022). Allostasis, Action, and Affect in Depression: Insights from the Theory of Constructed Emotion. *Annual Review of Clinical Psychology*, *18*(1), 553–580. https://doi.org/10.1146/annurev-clinpsy-081219-115627

The Lancet Global Health. (2020). Mental health matters. *The Lancet Global Health*, *8*(11), e1352. https://doi.org/10.1016/S2214-109X(20)30432-0

Van den Heuvel, M. P., & Sporns, O. (2013). An anatomical substrate for integration among functional networks in human cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *33*(36), 14489–14500. https://doi.org/10.1523/JNEUROSCI.2128-13.2013

Westlin, C., Theriault, J. E., Katsumi, Y., Nieto-Castanon, A., Kucyi, A., Ruf, S. F., Brown, S. M., Pavel, M., Erdogmus, D., Brooks, D. H., Quigley, K. S., Whitfield-Gabrieli, S., & Barrett, L. F. (2023). Improving the study of brain-behavior relationships by revisiting basic assumptions. *Trends in Cognitive Sciences*, *27*(3), 246–257. https://doi.org/10.1016/j.tics.2022.12.015

World Health Organization. World mental health report: transforming mental health for all. Geneva: World Health Organization, 2022