

Computational Psychiatry of Emotions and the Promise of it: Where it May Lead us

Iliyan Ivanov^{1*} and Jeffrey Schwartz²

- 1 Icahn School of Medicine, Mount Sinai, New York, USA
- 2 University of California, Los Angeles, USA

Abstract

Computational models can be used to enhance the description of behavioral responses that arise from the interaction between the brain and its immediate environment, and mathematical approaches are increasingly being utilized in the attempt to understand how the brain processes sensory data to elicit various behavioral responses. Computational psychiatry is the application of computational models to understanding psychopathology, with the modelling of mood and anxiety disorders becoming an area of particular interest. Various approaches to constructing these models have been suggested. For example, some investigators have proposed that distinguishing between “mood” and “emotions” is a key to computational modeling in studying depression, while others have pointed out that clinical phenomena related to the experience of depression and anxiety may be best accounted for by computational approaches to narratives patients construct to attempt to “make sense” of their biological condition. The growth and dynamic interaction of these various approaches has become an area with increasing potential for clinical application.

This paper presents a broad overview of the existing literature on the use of computational modeling in mood and anxiety disorders, as well as recent publications outlining the application of other related translational methods in psychiatry. We also present some original ideas on how Bayesian computational approaches can help explain the effects of attentional focus on brain function, and the how this relates to the utility of computational modeling for enhancing treatment interventions.

Keywords: Behavioral responses; Clinical application; Computational psychiatry; Emotions

Received: October 24, 2018; **Accepted:** December 10, 2018; **Published:** December 24, 2018

Introduction

Computational psychiatry has emerged as a new approach of applying mathematical models to describe behaviors that emerge in result of the interaction between the brain’s neurobiology and its immediate environment. Computational models of cognition can infer the current state of the environment and describe responses to environmental stimuli via concepts like reinforcement learning and predictive coding. Reinforcement learning for instance describes how the brain can choose and value courses of actions according to their future value, while predictive coding explains how the brain uses sensory data to infer about the state of its environment by combining

environmental signals with prior beliefs, each weighted according to their precision. As these concepts are applicable to normative behaviors they also can be applied to detect deviations from the normative range and to describe such “abnormal” behaviors that presumably correspond to symptoms of mental illness. Therefore computational models can capture behavioral difference between individuals diagnosed with a psychiatric condition and unaffected counterparts. In that sense computational psychiatry uses mechanistic descriptions of behaviors in order to enable the delineation of quantifying parameters that will be: 1) specific for particular cognitive functions 2) can be reliably measured and 3) will help assess how individuals with psychiatric conditions may differ from unaffected counterparts.

***Corresponding author:** Iliyan Ivanov

✉ iliyani@yahoo.com;
iliyan.ivanov@mssm.edu

Associate Professor of Child and Adolescent Psychiatry, President of the New York Council on Child and Adolescent Psychiatry, Icahn School of Medicine at Mount Sinai, New York, USA.

Tel: (646) 3876080

Citation: Ivanov I, Schwartz J (2018) Computational Psychiatry of Emotions and the Promise of it: Where it May Lead us. Acta Psychopathol Vol.4 No.6:24

Prediction Errors and Bayesian Inference

While for most part computational models deal with concepts like prediction error and learning, which in turn relate to cognitive processes like attention, working memory, decision making and alike it has been recognized that there are relations between computational models and emotions. A recent paper by Seth and Friston [1] has discussed the applicability of computational models in relation to emotional states. Their thesis is based on the concept is the so called “Bayesian brain” and related concepts like “approximate”, “perceptual” and “active” inference. For instance, *approximate inference* is typically used in ‘predictive coding’, which has been extensively studied in the field of machine learning. The brain is viewed as a system capable of predicting sensory data, which in turn is compared to the observed effects, and the difference between these constitutes the so called ‘prediction errors’. The prediction errors are further used to update the prior beliefs and assumptions and to transform them into expectations about possible outcomes. In order to minimize prediction errors in the future, the brain utilizes “perceptual” and ‘active” inference. Specifically, the goal of perceptual inference is to update the existing internal model of the world, while in active inference the goal is to change the sampling of the environment in order to better inform and adjust prior beliefs about the world. Therefore, one can either change empirical prior beliefs (through perception) to make predictions that are closer to sensations or one can change sensations (through increased sampling and observations of the environment) to make them closer to predictions (later in the paper we will discuss this latter notion in respect to future directions for integrating the Bayesian and quantum brain). For the purpose of their main objective Seth and Friston introduce the concept of “interoception”, which refers to the processes of detection and integration of autonomic signals “from within” the organism originating in the different biological systems (e.g. hormonal, immunologic etc). These interoceptive signals are thought to be essential in the development of emotional and embodied selfhood experiences. What is relevant to our topic is that these perceptions, described as “interoceptive inference”, can also be seen as parallel (or a modality-specific instance of) to “active inference” – in other words sensory systems produce ascending signals (e.g. prediction errors) that are compared with descending predictions across a hierarchy of perceptual processing. As these ideas are based on number of existing theories (cybernetic theories of regulation [2], perceptual control theory [3]) importantly this perspective emphasizes the anticipatory control-oriented nature of interoceptive inference [4] suggesting that predictive models that are used to describe behaviors (e.g. active inference) can also be applied to the interoceptive domain. The practical implications of these ideas are highlighted by their application to a variety of clinical conditions in which atypical interoceptive inference may play important roles in symptom development and expression.

Model-Based (MB) and Model-Free (MF) Cognition

For instance, a recent review by Adams et al. [5] discusses the relevance of using computational modeling to better understand the emergence of depressive psychopathology. By utilizing Bayesian Decision Theory (BDT), which encompasses reinforcement learning (RL), a comprehensive account of the structure of evaluative choice can be used to organize thinking about the breakdown of decision making in psychiatric disorders. BDT specifies that actions that are chosen would be expected to lead to maximal future utility. The field of reinforcement learning (RL) has delineated two fundamentally different ways in which past experience can be used to estimate and predict future positive and negative reinforcement (e.g., rewards and punishments), termed model-based (MB) and model-free (MF) cognition. MB cognition, which is essentially a form of goal-directed cognition, takes the form of analyzing the causes and consequences of an undetermined and variably sized array of potential events and actions related to them. Given a particular situation, the model can be searched, and the consequences of various behaviors deduced, even if they involve events that have not been previously experienced. Since this involves probing the model to simulate or infer future possibilities, it can have high computational costs. Alternatively, in MF cognition, which can be understood as a form of habitual learning, a model containing information about what is likely to transpire if a particular action is performed is not consulted. Instead, actions are chosen merely on the basis of how much and what kind of reinforcement has been obtained in the past when a certain state has been experienced or a particular action taken.

This will potentially reveal a set of values that accurately reflect future outcomes, to the extent that these outcomes have actually followed past choices.

This approach to information analysis has been used to formulate computational descriptions of the ways in which individuals suffering from depression experience reality and make decisions based on their approach to weighing various options, which reflects the manner in which they value past, present and future choices. Interestingly, the changes in value (or utility) ascribed to primary sensory stimuli do not seem to vary consistently in depressed subjects -- depression does not moderate responses to sweet tastes or systematically change pain sensitivity. MF types of habitual learning via reinforcement analysis do not adequately account for anhedonia, the reduced interest and/or pleasure that comprise a core symptom of depression. MB analyses of valuation account considerably better for describing the changes in valuation seen in the anhedonia that is characteristic of depression.

Specifically, MB learning seems to align with the phenomenon of learned helplessness, a particularly well investigated animal model of depression. In this model healthy animals exposed to uncontrollable stressors come to show a variety of depression-like behavioral anomalies. From a Bayesian perspective these effects can arise from prior beliefs about the achievability of

desirable outcomes. These prior beliefs would act within the goal-directed MB valuation system in ways that can account for diminished valuations seen in anhedonic depressive phenomena.

This could be reflected in a deficit in appropriate inhibition of the repetitive processing of aversive stimuli (sometimes called “aversive stimuli hijacking”) known to occur in depression. This is a potential cause of the repetitive negative thought pattern typical of rumination, a well-known aspect of the depressive cognitive processing.

A Bayesian Approach to Emotion and Mood

Others have adopted a different strategy by investigating the statistical relationship between “emotions,” which are elicited by discrete events, and “moods,” which last longer and tend to reflect the cumulative impact of a series of emotional responses. Depending on the degree of “mood stability,” mood can thus tend to either change gradually, as an aggregated effect of multiple preceding outcomes, or change more abruptly in response to outcome signals. People whose mood tends to be less stable will be significantly affected by unexpected changes in emotional state. In contrast, those with more stable moods will tend not to show suitable feedback interactions between unexpected outcomes and their associated moods. This can be experimentally studied, and computationally modeled, by investigating how unexpected changes in emotional state affect the valuation of future rewards. In active inference, rewards are defined in terms of prior preferences; i.e., beliefs that certain (rewarding) outcomes will result from action. The action that fulfills these prior beliefs is then selected to maximize expected reward.

Eldar and Niv [6] reported outcomes of experiments of this kind. Emotional state was manipulated by using a game of chance in which subjects either won or lost a relatively large sum of money. In subjects identified as being less emotionally stable, winning or losing the game significantly influenced both self-reported happiness and the effect of subsequent rewards on mood. Furthermore, striatal BOLD response to reward measured by fMRI on subsequent trials was also significantly affected. In contrast, subjects with more stable mood showed no such interaction between unexpected outcomes, their associated mood, and the valuation of future rewards. Thus, subjects characterized as having “mood instability” showed greater effects of emotional outcomes on mood, both in terms of behavioral and neurobiological responses. In contrast, subjects with more stable moods did not show significant effects of chance outcomes on future perceptions and neurobiological responses.

The concept of mood stability is thus being investigated in a computational model, in which unexpected outcomes affect emotional state and emotional state affects perception of future outcomes. This computational approach to emotion has clear face validity as a model of mood, and thus has potential clinical relevance. Although preliminary, these results suggest interactions between mood and learning that can potentially help

identify individuals with high vs. low mood stability [6,7]. Taken together, these ideas outline possible mechanisms for relations between computational models of information processing and learning, on one hand, and emotional states on the other. Still, while these models point to possible mechanisms, they do not explain how these bidirectional relations may result in the development of depressive illness.

These developments are exciting and seem to point to new ways of applying mathematical modeling to variety of behaviors that are thought to reflect both cognitive and emotional processing. As these developments appear to open opportunities for new lines of research it should be noted that the overall approach of “modeling” of human behavior faces some challenges. Possible limitations of the “modeling” approach in psychiatric disorders are discussed more recently by others [8]. In particular, in their review LeDoux and Pine point out that the inferred connections between the structures and the functions of brain circuits in the limbic system and the processing of emotional signals is still insufficient to explain the complexity of the human experience in relation to anxiety disorders. Further, the authors propose a two-system frame work that examines the interaction between cortical and subcortical brain systems in order to incorporate concepts like human conscious experience and the use of language within the larger framework of human psychopathology. This properly represents an acknowledgement of the fact that despite similarities in biological structure (e.g. brain neurocircuits) that translate into similarities and overlaps in behavioral patterns, which in turn can be mathematically modeled, it is premature to assume a direct link between computational models that describe behaviors and the true experience of mental illness in the human case. When the former can certainly define parameters for behavioral measures that can reflect symptoms, the latter seems to relate to the holistic conscious experience of having the illness of “depression” or “anxiety” or alike. In another publication LeDoux and Hofmann [9] argue that objective behavioral and physiological manifestations of emotions constitute only indirect indicators of the inner conscious experience of emotions and further suggest that the most direct way to assess conscious emotional feelings is through verbal self-report. The authors point out certain methodological barriers to studies of conscious emotional experience. First is the barrier of determining if non-verbal animals have emotional experiences. The counter argument is that since brain circuits underpinning emotions are highly conserved across species, studies of animals have an important role in understanding how emotions are expressed and regulated in the brain. This approach has been linked to a second problem, namely that conflation of circuits that directly give rise to conscious emotional experiences with circuits that indirectly influences conscious feelings has hampered progress in efforts to understand emotions, and also to understand and to develop treatments for emotional disorders. Furthermore, they argue that subjective reports – in their case of fear and anxiety - should not be viewed as unreliable proxy measures of fear and anxiety, but rather as the best and only measures accessible for clinicians to assess clinical cases.

The Role of Language and Conscious Experiences

From the view point of a clinician it is easy to attest that individuals with a major psychiatric illness create a personal “narrative” about “what kind of things are wrong with me” – in essence to create internalized “models” of expected outcomes/consequences of own behaviors (i.e. “nothing will work out”, “I will fail no matter what I do and how hard I try”) and how those beliefs reflect on one’s concepts of self-worth and self-value. Emotions associated with such personal models or concepts can in turn be conveyed exclusively via verbal reports. In relation to the discussed above feedback mechanisms between types of learning and “mood states” we suggest that for those to be relevant to the development of mental illness it is important to account for these internalized narratives about the relations between expectations/predictions, the behavioral choices related to those expectations and then the conscious interpretation of possible mismatched outcomes (e.g. “I try so hard but nothing works”). In those lines we strongly agree with the notion that conscious language-based narratives (or models) play a crucial role in the development of permanent emotional states that can become the illness of depression. As such language-based narratives emerge they also enter the human conscious experience. In other words as the feedback loops and model free/model based learning certainly contribute to the development of symptom of depression and anxiety, the internal narrative will serve to role of combining these symptoms into holistic conscious experience that will also influence the complex human behavior. Such proposition is in line with the two-system framework [8] in sense that the alterations in the information processing in the limbic system might be more directly linked to the development of neurovegetative symptoms of low energy, poor concentration, low motivation and alike (which often can be assigned to well defined neuronal circuits) whereas the contribution of the higher order cortical system will entail the development and internalization of personal narratives (or models) that will enter person’s conscious experience and create internal representations of low self-worth, hopelessness and helplessness that are not easy to map onto particular networks but rather represent “mind” constructs that will have the most profound effect on a person’s ability to function. In other words the former phenomenon will be linked to “feeling depressed” and latter to “having the illness of depression”.

Why is it that the role of language and conscious experience would be deemed so important in relation to the experience of emotions and emotional disorders? And why these issues would be relevant to computational approaches in psychiatry? First, we agree with the premise from DLoux and Pine that “Whatever degree of self-awareness and conscious experience is possible without language, it is clear that language changes the self-and-consciousness game in the brain”. They go on to suggest “... it is problematic to draw inferences about conscious experiences, and thus problematic to use words like fear and anxiety, when describing animal behavior and physiology. Moreover, because, as noted, defensive behaviors and physiological responding are not

reliable indicators of subjective fear in humans, such responses in animals cannot be used to identify circuits responsible for feelings of fear or anxiety in animals or humans”. As noted above, DeLoux and Hofmann [9] stress out the importance of verbal reports as possibly the only method to adequately ascertain the human aspect of constructs of fear and anxiety – a notion that is very likely true for other psychiatric disorders. In short, the role of language as a vehicle to assess the conscious experience seems indispensable in the process of assessing psychiatric conditions in humans. As animal models for such conditions clearly have no verbal communication components, the overall process of modeling faces a significant shortcoming. In that sense the relevance to computational modeling seems rather clear – as long as computational models are designed to describe behavioral measures with little attention to verbal reports they are likely to capture selected aspects of the syndrome but leave out important information. Of course, the use of verbal self-reports is well established in human research, however, these for most part are constructed on the base of questions that can be answered in binary “yes” or “no” fashion or ratings on Likert scale (i.e. the person will indicate the severity of their anxiety or depression in the range from 1 to 10). Such reports are rather different from a personal narrative that will provide a clinician with the information needed to design an individualized plan for treatment. That said there are many possible ways to incorporate verbal self-reports in computational methods and it seems imperative that such methods will be further developed. This is in line with a suggestion from LeDoux and Hofmann that “... future research re-examine the relationship between subjective report, behaviors, and physiological responses of fear and anxiety in order to explain the reasons for the dissociation between these measures [9].”

Second, and probably more importantly, understanding the role of language and conscious experiences in the complexity for mental illness is essential for its treatment. Often the goal of treatment is two-fold: 1) to reduce the severity of emotional distress and 2) to change types of behaviors that may perpetuate the symptoms of the illness. Psychiatric treatments in general consist of psychopharmacology and psychotherapy. While medications have shown efficacy in reducing the severity of certain symptoms – less intense anxiety, less frequent depressive episodes, improved energy and sleep - medication administration by itself rarely (if ever) will automatically results in changes of the overall behaviors that often are characteristic of any particular disorder. Moreover, medication administration is not likely to change the personal narrative or the experience of the disease towards healthier outlooks on one’s experiences in life. Clearly such types of changes are primarily accomplished through psychotherapy interventions that are all conducted via the use of language with the goal that new and different ideas will arise in the patients’ conscious experience and will result in efforts to change “how one does things”. Psychotherapy has been successful in both alleviating symptoms of depression and anxiety and producing changes in behaviors that can further lead to decreased suffering. Here we suggest that computational methods can be applied to different aspects of psychotherapy treatments (**Figure 1**).

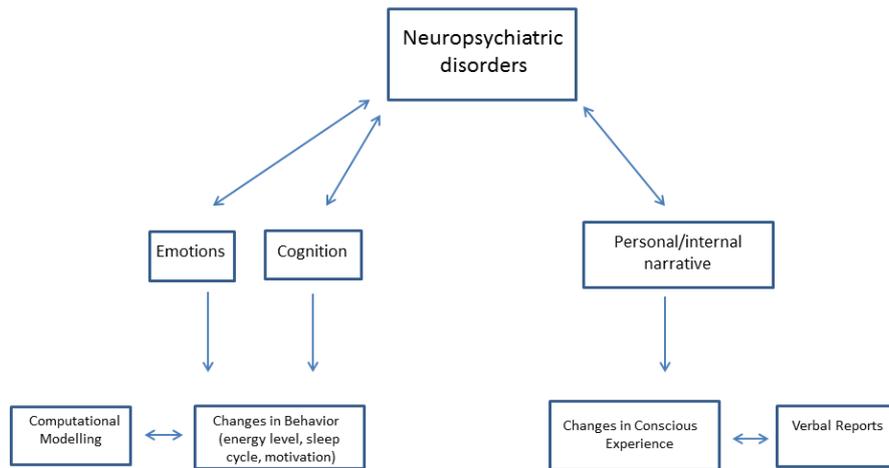


Figure 1 This figure illustrates a two system model of neuropsychiatric disorders where emotional and cognitive processes produce changes in motivation, attention, activity levels, sleep and appetite that in turn can be further assessed by computational modelling whereas the personal narrative influences one's conscious experience which can be accessed via verbal reports.

Mindfulness, Quantum Zeno Effect (QZE) and the Bayesian Brain

In the recent years techniques related to meditation and mindfulness have been increasingly studied in relation to their efficacy in addressing emotional distress. Studies have shown that mindfulness decreases the severity of symptoms of anxiety and depression in variety of psychiatric conditions [10,11] including substance use [12,13] and Posttraumatic Stress Disorder [14,15]. In addition, it has been demonstrated that mindfulness-based practices have effects on brain structure and function [16,17] and particularly may influence functional connectivity between the lateral prefrontal cortices implicated in self-monitoring and cognitive control, and medial prefrontal and other default-mode network areas associated with mind-wandering and self-referential thoughts [18-20]. Further, structural MRI has shown significantly increased gray matter density and cortical thickness in lateral temporal cortices, anterior cingulate, and hippocampus in individuals who had intensive training in the memorization and recitation of ancient scriptural texts, techniques that are traditionally used in the practice of meditation [21]. In short, there is significant emerging evidence suggesting that techniques utilized in the practice of mindfulness and meditation relieve emotional distress, reduce intensity of depression and anxiety, and are associated with changes in brain structure and function.

Prior work by one of us, in collaboration with physicist Henry Stapp [22], suggests a possible physical mechanism for sustained mindfulness to produce a stabilizing effect on brain function. Specifically, it has been theorized that focused attention can produce stabilizing effects on functional connectivity of neural circuitry via a well-recognized phenomenon of quantum mechanics called quantum Zeno effect [22]. The quantum Zeno effect (QZE), sometimes called the "Watch Dog effect", is a well-described feature of quantum mechanics which states that the rate of change of the system will be markedly decreased

by frequent measurement. This effect, named for the Greek philosopher Zeno of Elea, was brought into prominence in 1977 by the physicists Misra & Sudarshan [23] and has been experimentally confirmed many times [24,25]. It gives a name to the fact that repeated and closely spaced observational acts can effectively hold a prior observation in place for an extended time-interval that depends upon the rapidity at which subsequent observational actions occur. According to a model developed to apply this quantum principle to the effects of acts of directed attention on brain function [22], this rapidity of observation can be controlled by the amount of effort applied in successive acts of repeatedly focusing attention. This model is particularly well adapted to describe the effects of mindfully directed attention on functional brain activity.

The link between the quantum zeno effect and functional integration in the brain is particularly prescient for predictive coding formulations of active inference (and by implication the free energy principle). This follows because in predictive coding, hierarchical inference involves reciprocal message passing between hierarchical levels. These messages correspond to precision-weighted prediction errors – that ascend from one level to the next – and descending predictions, which define the prediction errors. Crucially, the influence of prediction errors on (Bayesian) belief updating in higher levels is controlled by their precision. This means the brain should also be equipped with descending predictions of precision.

Psychologically, this may be the best candidate to explain attention and attentional states (such as mindfulness). Physiologically, precision weighting rests upon the neuronal excitability of populations encoding prediction errors (e.g., superficial pyramidal cells). If one subscribes to this computational architecture, then the quantum zeno effect – operating on descending predictions of precision – can only be manifest in terms of a change in the connectivity among hierarchical brain levels.

The importance of ‘precision engineered’ message passing in the brain also speaks to the notion of *mental action* in the following sense: to actively attend to – or select – particular sets of prediction errors, it is necessary to infer their precision in any given context. This capacity has been associated with mental action and, indeed, the phenomenal opacity of our perceptual experiences. On the present argument, one could argue that the quantum zeno effect, mediated by repeated mental actions, will reduce the rate of change of predicted precision and, in the interoceptive domain, underwrite mood stabilization.

Interestingly, this formulation is entirely consistent with accounts of affective psychopathology that, generally, call on a failure to attenuate low-level of sensory precision. In other words, many affective symptoms can be ascribed to an inability to attend away from interoceptive cues (e.g., elevated interoceptive sensitivity), which would be associated with instability of mood and inappropriate belief updating. In turn, this suggests that the stabilizing effect of any mental action may constitute a therapeutic approach to both attentional and affective states. Focused attention attained in the process of mindfulness can result in an increased frequency of monitoring (i.e. higher sampling rate) while at the same time minimizing incoming signals from immediate surroundings (e.g. meditative focusing on breathing and volitionally ignoring other stimuli) and in theory will optimize or stabilize information processing in brain networks. This optimization process can also contribute to

enhanced capacities in exercising processes of active inference and goal-directed-decision making [26,27] via increased abilities to monitor and process information relevant to current states of the world/body, potentially attainable states, and more accurate assessments regarding future potential goal states.

In short, the Bayesian brain attempts to minimize free energy by minimizing the difference between expectations and outcomes. This is accomplished by changing the rate of sampling of the environment and/or changing the internal representation of the outside stimulus. Close continuous monitoring of a stimulus will more precisely detect small changes, and will constantly update the internal representation of it so that the difference between expectation and outcome will be minimized, trending toward zero. In a related manner, the quantum zeno effect states that the act of sustained observation will stabilize and effectively hold a prior observation in place. The practice of mindfulness and focused attention may achieve and actualize effects of this kind in brain circuitry – for instance focusing attention on a limited set of stimuli (one’s breathing) and at the same time on a limited set of internal representations (one’s awareness of own breathing) will in theory minimize the difference between expectation and outcome. A potentially illuminating analogy to this effect of observation would be watching a segment of a movie in slow motion – one can redirect attention to any particular detail and register very small changes while simultaneously updating their representation of that particular detail, so that the amount of novelty and surprise will be diminished to a minimum (**Figure 2**).

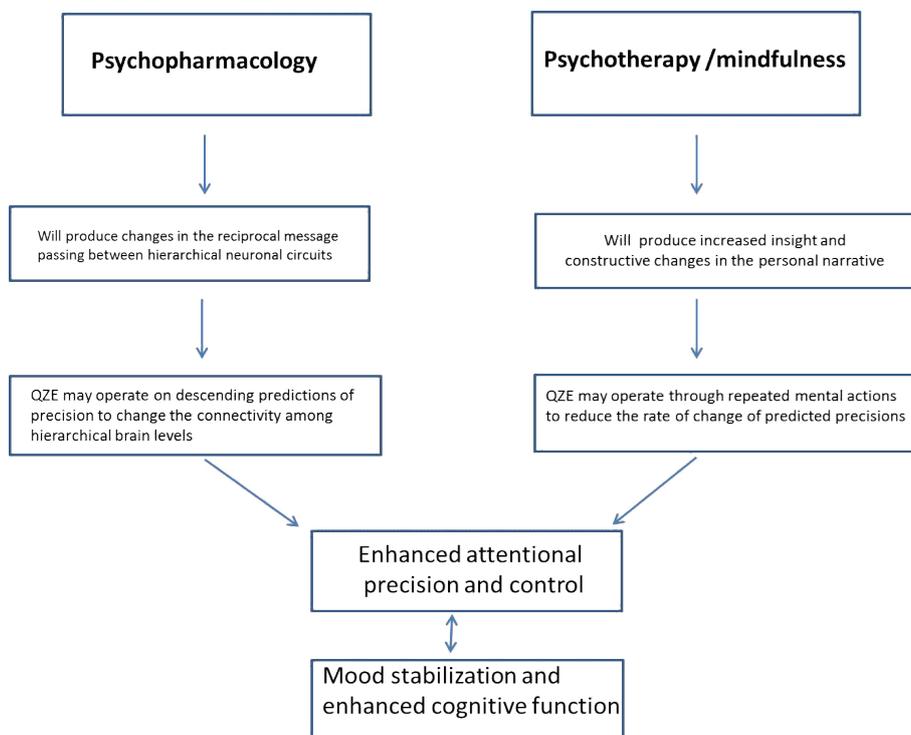


Figure 2 Schematic presentation of the effects of different interventions (i.e. pharmacotherapy and psychotherapy) that produce changes in the messages passing in neuronal circuits, and in individual’s personal narrative, respectively. It is further suggested that these influences may be modulated by quantum zeno effects (QZE) that are operational at the level of predicted precisions, which in turn may lead to mood stabilization.

System-Environment Interactions

The other aspect of treatments relevant to computational principles is related to the effects of changes in individual's behavior to its immediate environment. These interaction effects between a system and environment have been described in computational terms also through the free energy principle (i.e., active Bayesian inference). In general several types of system- environment interactions have been described including increased sampling of environment to update the information and also adjustment of internal representations – both leading to a minimization of free energy or, under some simplifying assumptions, (precision weighted) prediction errors. In terms of human psychology these processes are relevant phenomena like reframing and relabeling that have been successfully used in psychotherapy to alleviate symptoms of anxiety and depression.

The human brain can also execute additional options like action upon the environment in order to change environmental factors. The computational links between free energy and action undertaken by biological systems are discussed in detail by Peters et al. [26]. Even further, computational methods have examined processes that precede and are necessary for the emergence of active action on the environment. For instance, goal-directed decision making, which is essential for setting up series of actions that can lead to an instant outcome, can be understood in terms of active inference [27,28]. This idea suggests that the selection of behavioral strategies can be treated as an inference problem when one considers three kinds of probability distributions; 1) current states of the system and its environment, 2) states that can be reached, i.e., attainable states, and 3) goal states that the agent believes he/she should occupy.

Strategy selection occurs under the prior belief that it will minimize the difference between the probability distribution over 'attainable states' (given beliefs about the current state) and the probability distribution over 'goal states' and that behavioral choices are based upon beliefs about alternative strategies, where the most likely strategy minimizes the difference between attainable and desired outcomes (repertoire versus goal) [27,28]. Some of us have recently proposed a biological model of deliberation (i.e. choice of action without action) that is also based on the principles of inference [29]. The relevance of these concepts of goal directed inference and deliberation is that these seem crucial when an organism can minimize free energy (i.e., precision weighted prediction errors) mainly by introducing actively changes in its immediate environment. Therefore, it is reasonable to suggest that active inference and deliberation are integral for the social functioning of an individual which in turn is deemed an essential criterion for diagnosing a psychiatric disorder. It is also reasonable to suggest that creating narratives that will reflect general beliefs about the state of the world will strongly influence these processes and will in turn influence one's ability to function socially. From the perspective of integrating free energy principles and the clinical view of psychiatric disorders as impaired social functioning the development of certain psychiatric symptoms can be seen as an adaptive response to altered computation- anhedonia might be an attempt to

minimize environmental inputs that are distressing and may continuously influence the experience of feeling unhappy or bad. Alternatively attempts to change the immediate environment that may stem from flawed perception and interpretations may present as disorganized, not reality based behaviors. When the conscious experience is not necessarily required for these types of responses per se it is reasonable to suggest that the cornerstone of psychiatric diagnosing - impaired functioning - does not simply refer to having depressive/anxiety symptoms but to the profound change in complex social behaviors that require conscious guidance.

In summary, this literature provides meaningful support for the hypothesis that computational models based on concepts like active inference may be able to reliably and accurately describe individual behavior in a clinically relevant manner. These behaviors can interface and interact with emotional experiences (feeling happy or sad) and thus create feedback loops that have dynamical effects on both the behaviors and the emotions. Those feedback processes provide data that can be used to assess the differences between expected and actual rewards (e.g. prediction errors). For example, situations with elevated stress are associated with prediction errors that are found to be encoded with higher precision, possibly to optimize the system's response in conditions when accurate responding might be crucial for survival [26].

While examples such as this are already within the purview of computational psychiatry, it is also possible that computational principles are integral to more complex behaviors, such as goal directed decision making and deliberation - and these latter functions are related to the person's ability to function in relation to other persons (social functioning). For instance, in individuals with mental illness constitutional alterations in computational abilities may result in aberrant encoding of precision, and these abnormalities can create the background for the development of internal narratives that are excessively specific to the individual's own experience, with a resulting decrement in social functioning. This might be essential to transforming "abnormalities" of computation into chronic allostatic feedback imbalances that can become manifested as mental illness. This development of internal states/models that idiosyncratically reflect one's perception of the inner environment can be understood as an essential part of the development of mental disorders. Language use and misuse may play a crucial role in that process. It is becoming increasingly possible to use analytic methods to assess information processing related to language, and investigate whether any purported abnormalities may indeed differentiate between affected and unaffected individuals.

Developments of this kind may enhance our ability to identify and diagnose mood and anxiety disorders, and potentially guide the development of new treatments. In that context, we hypothesize that there might be a dynamical interactive connection between the Bayesian brain and the Quantum Zeno Effect that both enhances our understanding of the therapeutic utility of meditation and mindfulness as interventions in emotional disturbances and disorders, and amplifies the potential synergies

of these interventions with psychopharmacological treatments. This last point is of crucial importance to any future research that aims to not simply “understand” mental illness but to possibly develop more effective treatments.

References

- 1 Seth AK, Friston KJ (2017) Active interoceptive inference and the emotional brain. *Philos Trans R Soc Lond B Biol Sci* 371: 20160007.
- 2 Conant R, Ashby WR (1970) Every good regulator of a system must be a model of that system. *Int J Syst Sci* 1: 89-97.
- 3 Srinivasan MV, Laughlin SB, Dubs A (1982) Predictive coding: a fresh view of inhibition in the retina. *Proc R Soc Lond B Biol Sci* 216: 427-459.
- 4 Seth A, Critchley HD (2013) Extending predictive processing to the body: emotion as interoceptive inference. *Behav Brain Sci* 36: 227-228.
- 5 Adams RA, Huys QJ, Roiser JP (2015) Computational psychiatry: towards a mathematically informed understanding of mental illness. *J Neurol Neurosurg Psychiatry* 87: 53-63.
- 6 Eldar E, Rutledge RB, Dolan RJ, Niv Y (2016) Mood as representation of momentum. *Trends Cogn Sci* 20: 15-24.
- 7 Huang H, Thompson W, Paulus MP (2017) Computational dysfunctions in anxiety: failure to differentiate signal from noise. *Biol Psychiatry* 82: 440-446.
- 8 LeDoux JE, Pine DS (2016) Using neuroscience to help understand fear and anxiety: a two-system framework. *Am J Psychiatry* 173: 1083-1093.
- 9 LeDoux JE, Hofmann SG (2018) The subjective experience of emotion: a fearful view. *Curr Opin Behav Sci* 19: 67-72.
- 10 Goyal M, Singh S, Sibinga EMS, Gould NF, Rowland-Seymour A, et al. (2014) Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. *JAMA Intern Med* 174: 357-368.
- 11 Groves P (2016) Mindfulness in psychiatry - where are we now? *BJPsych Bull* 40: 289-292.
- 12 Bowen S, Witkiewitz K, Clifasefi SL, Grow J, Chawla N, et al. (2014) Relative efficacy of mindfulness-based relapse prevention, standard relapse prevention, and treatment as usual for substance use disorders: a randomized clinical trial. *JAMA Psychiatry* 71: 547-556.
- 13 Glasner S, Mooney LJ, Ang A, Garneau HC, Hartwell E (2017) Mindfulness-based relapse prevention for stimulant dependent adults: a pilot RCT. *Mindfulness (NY)* 8: 126-135.
- 14 Kearney DJ, McDermott K, Malte C, Martinez M, Simpson TL (2013) Effects of participation in a mindfulness program for veterans with posttraumatic stress disorder: a randomized controlled pilot study. *J Clin Psychol* 69: 14-27.
- 15 Smith BW, Ortiz JA, Steffen LE, Tooley EM, Wiggins KT, et al. (2011) Mindfulness is associated with fewer PTSD symptoms, depressive

Acknowledgement

Karl J. Friston provided critical assistance in articulating the arguments in this article.

- symptoms, physical symptoms, and alcohol problems in urban firefighters. *J Consult Clin Psychol* 79: 613-617.
- 16 Pickut BA, Van Hecke W, Kerckhofs E, Mariëne P, Vanneste S, et al. (2013) Mindfulness based intervention in Parkinson's disease leads to structural brain change on MRI: a randomized controlled longitudinal trial. *Clin Neurol Neurosurg* 115: 2419-2425.
- 17 Tang YY, Hölzel BK, Posner MI (2015) The neuroscience of mindfulness meditation. *Nature Rev Neurosci* 16: 213-225.
- 18 Brewer JA, Worhunskya PD, Gray JR, Tang YY, Weber J, et al. (2011) Meditation experience is associated with differences in default mode network activity and connectivity. *Proc Natl Acad Sci USA* 108: 20254-20259.
- 19 Hasenkamp W, Barsalou LW (2012) Effects of meditation experience on functional connectivity of distributed brain networks. *Front Hum Neurosci* 6: 38.
- 20 Lin Y, Callahan CP, Moser JS (2018) A mind full of self: self-referential processing as a mechanism underlying the therapeutic effects of mindfulness training on internalizing disorders. *Neurosci Biobehav Rev* 92: 172-186.
- 21 Hartzell JF, Davis B, Melcher D, Miceli G, Jovicich J, et al. (2016) Brains of verbal memory specialists show anatomical differences in language, memory and visual systems. *Neuroimage* 131: 181-192.
- 22 Schwartz JM, Stapp HP, Beauregard M (2005) Quantum physics in neuroscience and psychology: a neurophysical model of mind-brain interaction. *Philos Trans R Soc Lond B Biol Sci* 360: 1309-1327.
- 23 Misra B, Sudarshan ECG (1977) The Zeno's paradox in quantum theory. *J Math Phys* 18: 756-763.
- 24 <https://phys.org/news/2015-10-zeno-effect-verifiedatoms-wont.html>
- 25 Patil YS, Chakram S, Vengalattore M (2015) Measurement-Induced Localization of an Ultracold Lattice Gas. *Phys Rev Lett* 115: 140402.
- 26 Peters A, McEwen BS, Friston K (2017) Uncertainty and stress: why it causes diseases and how it is mastered by the brain. *Prog Neurobiol* 156: 164-188.
- 27 Friston K, Schwartenbeck P, Fitzgerald T, Moutoussis M, Behrens T, et al. (2013) The anatomy of choice: active inference and agency. *Front Hum Neurosci* 7: 598.
- 28 Friston K, Schwartenbeck P, Fitzgerald T, Moutoussis M, Behrens T, et al. (2014) The anatomy of choice: dopamine and decision-making. *Philos Trans R Soc Lond B Biol Sci* 369: 20130481.
- 29 Ivanov I, Whiteside K (2017) Dyadic brain-a biological model for deliberative inference. *AIMS Neuroscience* 4: 169-188.