



OPEN ACCESS

EDITED BY
Povilas Karvelis,
University of Toronto, Canada

REVIEWED BY
Subutai Ahmad,
Numenta Inc., United States
Mircea R. Stan,
University of Virginia, United States

*CORRESPONDENCE
Mohamed A. Sherif
✉ Mohamed_sherif@brown.edu

†PRESENT ADDRESS
Joshua C. Brown,
Department of Psychiatry, Harvard Medical
School, McLean Hospital, Boston, MA,
United States

SPECIALTY SECTION
This article was submitted to
Computational Psychiatry,
a section of the journal
Frontiers in Psychiatry

RECEIVED 23 June 2022
ACCEPTED 16 January 2023
PUBLISHED 23 February 2023

CITATION
Sherif MA, Khalil MZ, Shukla R, Brown JC and
Carpenter LL (2023) Synapses, predictions,
and prediction errors: A neocortical
computational study of MDD using
the temporal memory algorithm of HTM.
Front. Psychiatry 14:976921.
doi: 10.3389/fpsy.2023.976921

COPYRIGHT
© 2023 Sherif, Khalil, Shukla, Brown and
Carpenter. This is an open-access article
distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

Synapses, predictions, and prediction errors: A neocortical computational study of MDD using the temporal memory algorithm of HTM

Mohamed A. Sherif^{1*}, Mostafa Z. Khalil², Rammohan Shukla³,
Joshua C. Brown^{4†} and Linda L. Carpenter⁴

¹Lifespan Physician Group, Department of Psychiatry and Human Behavior, The Warren Alpert Medical School of Brown University, Carney Institute for Brain Science, Norman Prince Neurosciences Institute, Providence, RI, United States, ²Department of Psychiatry and Behavioral Health, Penn State Milton S. Hershey Medical Center, Penn State College of Medicine, Hershey, PA, United States, ³Department of Neurosciences, The University of Toledo College of Medicine and Life Sciences, Toledo, OH, United States, ⁴Department of Psychiatry and Human Behavior, The Warren Alpert Medical School of Brown University, Butler Hospital, Providence, RI, United States

Introduction: Synapses and spines play a significant role in major depressive disorder (MDD) pathophysiology, recently highlighted by the rapid antidepressant effect of ketamine and psilocybin. According to the Bayesian brain and interoception perspectives, MDD is formalized as being stuck in affective states constantly predicting negative energy balance. To understand how spines and synapses relate to the predictive function of the neocortex and thus to symptoms, we used the temporal memory (TM), an unsupervised machine-learning algorithm. TM models a single neocortical layer, learns in real-time, and extracts and predicts temporal sequences. TM exhibits neocortical biological features such as sparse firing and continuous online learning using local Hebbian-learning rules.

Methods: We trained a TM model on random sequences of upper-case alphabetical letters, representing sequences of affective states. To model depression, we progressively destroyed synapses in the TM model and examined how that affected the predictive capacity of the network. We found that the number of predictions decreased non-linearly.

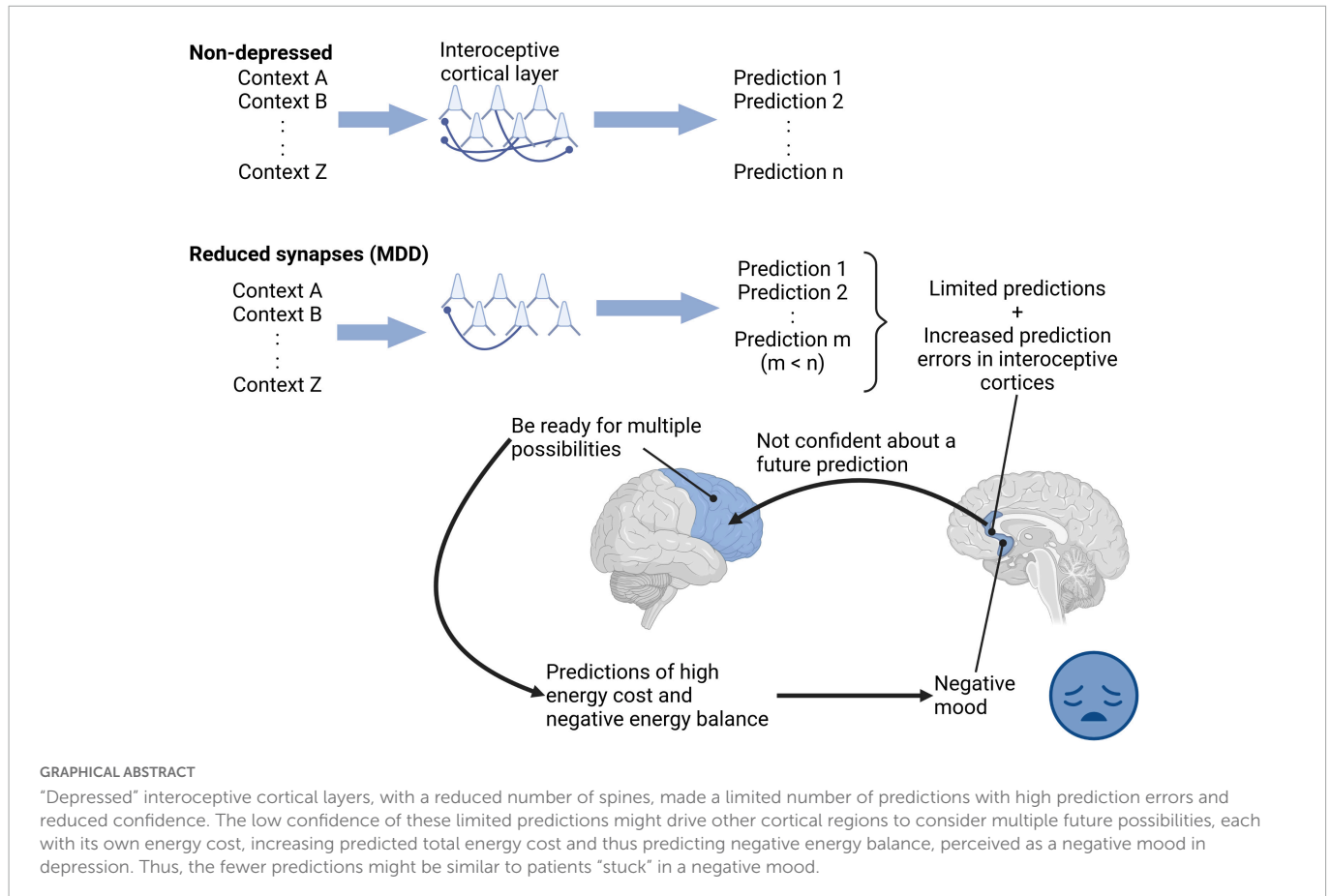
Results: Destroying 50% of the synapses slightly reduced the number of predictions, followed by a marked drop with further destruction. However, reducing the synapses by 25% distinctly dropped the confidence in the predictions. Therefore, even though the network was making accurate predictions, the network was no longer confident about these predictions.

Discussion: These findings explain how interoceptive cortices could be stuck in limited affective states with high prediction error. Connecting ketamine and psilocybin's proposed mechanism of action to depression pathophysiology, the growth of new synapses would allow representing more futuristic predictions with higher confidence. To our knowledge, this is the first study to use the TM model to

connect changes happening at synaptic levels to the Bayesian formulation of psychiatric symptomatology. Linking neurobiological abnormalities to symptoms will allow us to understand the mechanisms of treatments and possibly, develop new ones.

KEYWORDS

MDD, predictions, prediction error, ketamine, psilocybin, hierarchical temporal memory (HTM), confidence



Introduction

Synapses and spines play an integral role in the pathophysiology of major depressive disorder (MDD), a severely disabling disease that affects many aspects of life (1–3). Postmortem evidence shows a reduction in spine density (4) and decreased synapse numbers in the dorsolateral prefrontal cortex (DLPFC) in patients diagnosed with MDD (5). Numerous lines of pre-clinical evidence also highlight this relationship. Paired associative stimulation experiments using transcranial magnetic stimulation (TMS) showed impaired plasticity in patients struggling with MDD, suggesting abnormal synaptic functioning (6). In chronic stress animal models of MDD, dendritic spines of layer 2/3 in the medial prefrontal cortex are lost (7–9). Chronic imipramine or fluoxetine restored spine loss in both medial prefrontal cortex layer 2/3 pyramidal neurons and hippocampal CA3 in animal models of chronic stress (10). Likewise, a single session of repetitive TMS in hippocampal CA1 slice cultures increased spine size and functional insertion of the GluA1 subtype of AMPA receptors consistent with long-term potentiation (LTP) of excitatory synapses (11). Cambiaghi

et al. (12) showed increased total spine density in apical and basal dendrites of pyramidal neurons layer 2/3, with a pre-dominance of thin spines following 5 days of high frequency repetitive TMS applied to motor cortex. The application of repetitive TMS also increased dendritic complexity. Similar effects on spine maturation have been reported with electroconvulsive therapy in animal models (13, 14).

These same cellular mechanisms are suggested by results of cortical physiology experiments showing LTP-like changes in healthy humans after a single session of repetitive TMS delivered to motor cortex in combination with pharmacologic n-methyl-d-aspartate (NMDA) receptor agonism (15, 16). Such findings support the notion that effective antidepressant treatments such as repetitive TMS (17, 18), approved for depression in 2008 (19), may work through rescuing impaired synaptic plasticity. This notion received additional support from two recent clinical studies investigating the combination of stimulation with either placebo or the NMDA receptor agonist d-cycloserine; in the first, the active combination enhanced and normalized plasticity measures in the motor cortex of depressed patients (20, 21).

Single doses of ketamine (22–26) and psilocybin (27) result in rapid antidepressant effects in depressed patients that appear to be mediated by restoring synapses and spines. A single dose of ketamine increased the rate of spine formation in non-stressed animals (28). Following chronic stress, ketamine restored the number of neocortical synapses in animal models in layer 5 of the medial prefrontal cortex (29, 30). A single dose of ketamine also protected spine loss induced by chronic stress (31). The deleterious effects on spines were mediated by proteins like REDD1 that belong to the mTOR pathway in layer 5 apical dendrites, which increased with chronic stress. REDD1 was also elevated in the prefrontal cortex in patients diagnosed with MDD (32). An increase in spine densities suggests the formation of more synapses (33), and mediated synchronized firing of ensembles encoding behavioral manifestations affected by depression (34). Local infusion of ketamine into the infralimbic cortex of rats resulted in the reversal of depression-like behavior. This effect was also generated by optogenetic stimulation of layer five pyramidal neurons and resulted in similar increases in spine density in apical tufts (35). Related findings by Shao et al. showed that psilocybin increased dendritic spines on both apical and basal dendrites in the medial frontal cortex in a mouse model of depression, reversing the behavioral manifestations of depression (36). Together these clinical and preclinical findings converge to illustrate a critical role for synaptic function and number in the pathology of depression and in antidepressant treatment effects, but the direct and dynamic investigation of neuronal synapses in living clinical samples is not currently possible.

To bridge synaptic changes to symptomatology in MDD, it is helpful to consider the function of the central nervous system from an evolutionary perspective. The brain has evolved to optimize energy consumption as the body acts and grows, a phenomenon termed allostasis, which means stabilizing in the face of change (37–40). Instead of reacting to inputs, the neocortex develops an internal model of the world and of the body, and uses the model to predict input changes and update the model predictions based on mismatches with the inputs. This has been formulated into the Bayesian brain and predictive coding frameworks (41, 42).

In the motor or “action” realm, instead of minimizing prediction error by updating the model to match inputs, the body minimizes the prediction error by enacting the prediction. For example, before a person reaches out to grasp a cup of water, the motor cortex generates predictions about the action, such as “what would I experience when I am grasping that cup of water I see in front of me on the counter?” Relevant visual, tactile, and proprioceptive predictions are projected to the spinal cord. Spinal cord reflex arcs then “enact” or fulfill these predictions by moving the person’s hand and grasping the cup of water, thus, reducing prediction error. This is termed active inference (43).

Emotional states and their associated autonomic, hormonal, and immune system changes are hypothesized to reflect enactments of the brain’s predictions based on inputs from the inner milieu of the body, as well as from the external world (interoceptive predictions). Formulated in the Embodied Predictive Interoception Coding (EPIC) framework, mood has been considered a low-dimensional summary of the predictions related to these inputs (44). Interoceptive predictions occur in interoceptive limbic cortical areas, such as the anterior cingulate cortex (ACC) and insula,

the areas affected in chronic stress models of depression where ketamine was found to be helpful in animal models. Applying the interoceptive theory to understand states of clinical depression, Barrett et al. hypothesized that the brain is locked-in, i.e., stuck in a “metabolically-inefficient internal model of the body in the world” (37, 45), constantly predicting negative metabolic energy balance, i.e., that there will be less metabolic energy available in the future. According to the interoceptive theory, such predicted negative energy states will be experienced as unpleasant affective states. It will also result in experiences of reduced energy and consequently, decreased motivation. For example, hearing about an event that I enjoyed before, the pleasant affective state I experienced before is reflective of having overall more energy during the event. In an undepressed state, my brain would predict positive energy balance, i.e., more energy, during the event similar to my previous experience. But in a depressed state, my brain will be stuck predicting negative energy balance, i.e., that I will have less energy available during the event. This will be reflected as unpleasant affective state when I hear about an announcement for a future event, reducing my motivation to go. Such locked-in states can also manifest as ruminations, where the brain keeps coming back to the same thought content. These states could arise as a consequence of the abnormalities in interoceptive cortical areas that have been revealed by neuroimaging studies in depressed samples (46), and suggested as sites for DBS stimulation treatment for MDD (47). Indeed, TMS has better outcomes when *post hoc* modeling shows DLPFC stimulation has strong functional connectivity (negative correlation) with the ACC (48, 49). Taken further, personalized targeting to the ACC was one of several changes that produced the strongest TMS outcomes to date (50), and may become the standard for TMS therapy (51).

Machine learning algorithms constrained by neocortical neurobiology can connect synaptic atrophy to impaired predictions, a functional construct more proximal to symptoms. The neocortex is hypothesized to be a spatio-temporal pattern-predictive machine (52–54). The ability of the neocortex to learn from new information on the fly and adjust its predictions accordingly is an integral part of this process (55). This is referred to as “online learning” in the machine learning literature and is different from batch training, where an algorithm learns a training set, then is used on a different set. We used the temporal memory (TM) algorithm, a computer model of neocortical layers, to investigate the link between synaptic changes and prediction making. TM is one algorithm of the hierarchical temporal memory (HTM) framework. The underlying assumption of HTM is that different neocortical regions implement versions of the same cortical algorithm (56) for online unsupervised sequence learning and prediction (57, 58). Multiple features of the neocortex constrain HTM algorithms in ways unlike most of the current artificial neural networks (59). For example, the units or “neurons” in the HTM algorithms exist in one of three states: active, inactive, and predictive. Neurobiologically, the predictive state of a neuron maps to subthreshold depolarization of the soma by dendritic NMDA spikes, making it easier for the neuron to fire. NMDA spikes are triggered by the co-activation of synapses clustered together on a dendritic segment, indicating the detection of a specific input pattern (60). HTM exhibits online learning, where the algorithm continuously learns and predicts, rather than being trained in batches like artificial neural networks. In addition, the HTM algorithms learn

TABLE 1 Default parameter values of temporal memory (TM) model.

Number of minicolumns per letter	20
Column dimensions (number of columns)	520
Cells per column	32
Activation threshold (number of active synapses for a dendritic segment to be considered active)	13
Initial permanence	0.21
Connected permanence (permanence value above which a synapse is connected)	0.5
Min threshold	10
Max new synapse count	20
Permanence increment	0.1
Permanence decrement	0.1
Predicted segment decrement	0.0004
Max segments per cell	255
Max synapses per segment	255

by local unsupervised Hebbian-learning rules similar to synaptic plasticity that does not require error back propagation. HTM algorithms act on sparse-distributed representations (SDR) of data, which takes the stochastic nature of synaptic activation into account (57).

With a TM model representing a neocortical layer of neurons that learns and predicts sequences of inputs, this study explored a possible connection between having a reduced number of synapses (i.e., characteristic of MDD) and being locked in dysfunctional predictive states experienced as negative affective states. The TM model we use contains basal distal dendrites, which provide temporal context to the presented stimuli (57, 59). We ran simulations to examine how many predictions are generated given a particular context, and then examined how synaptic loss affected the total number and confidence of the predictions made.

Materials and methods

Inputs, temporal memory (TM) models, and training

The input patterns consisted of an array of ten randomly generated sequences of 11 upper-case letters (A–Z) (Figure 1). A TM model is composed of multiple minicolumns. Each upper-case letter was represented in the TM network by activating neurons in 20 consecutive minicolumns, resulting in a network with 520 minicolumns (20 minicolumns × 26 upper-case letters). Each minicolumn contained 32 neurons.

Using the Numenta Platform for Intelligent Computing (NuPIC) package (version 1.0.5) we generated five temporal memory (TM) models (parameters in Table 1) using five different randomization seeds. We then trained each model on five different versions of the input patterns, i.e., five different arrays, each of ten randomly generated sequences. A reset signal was applied at the end of each input sequence to reset the predictive neurons. The training steps are

explained in detail in the results section and in schematic Figure 2. On average, each TM model correctly learned each sequence of the 11 letters over ten repeated presentations. Therefore, for each of the five versions of the input pattern, we ran 1,100 iterations of simulations to ensure the model had learned the full input pattern. This generated five trained network models for each input pattern, resulting in 25 trained TM models. Next, we calculated the total number of future predictions a network can make by summing the number of predictions for each of the upper-case letters:

$$\text{Total predictions} = \sum_{\text{letter}} \text{number of predictions}_{\text{letter}}; \text{letter} = \{A \text{ to } Z\}$$

Destroying synapses

Using five randomization seeds that were different from the ones used to generate the TM models, we progressively removed fractions of synapses, i.e., we destroyed 25, then 50, then 60, then 70, and finally 80% of the total number of synapses. For each of the 25 networks (5 networks × 5 random variations of input sequences), this step created five different versions of “depressed” networks for each percentage of destroyed synapses, resulting in 125 versions of networks for each percentage of destroyed synapses. We then calculated the total number of predictions made at each level of synaptic destruction.

Prediction error and calculating confidence

As described in more detail in the Results section, when a minicolumn receives a predicted input, only the neurons in the predictive state fire, making them the winner cells. But when a minicolumn receives an unpredicted input, none of the cells in that minicolumn were in a predictive state, so all the neurons in the minicolumn fire (minicolumn burst) representing prediction error, and winner cells are selected at random from the active neurons. To estimate the prediction error, we used the ratio of winner cells to active cells to represent the confidence of the network about a prediction:

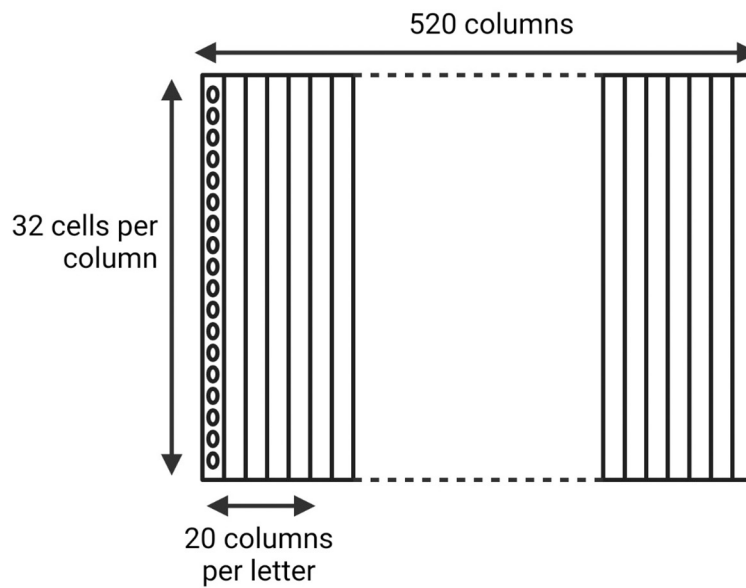
$$\text{Confidence} = \frac{\text{winner neurons}}{\text{active neurons}}$$

Lower confidence, usually due to a higher number of active cells when a minicolumn bursts (higher prediction error), means a higher mismatch occurred between the prediction and subsequent input. Higher confidence represents a better match between prediction and subsequent input. Confidence of a prediction is distinct from the total number of predictions a network can make which is a function of the size of the network and the history of inputs it was trained on.

Examining hyperparameters

We ran simulations where we varied the connected permanence, which is the value of spine permanence above which the spine is

Model: A single temporal memory (TM) layer.



Input: 10 sequences, each of 11 letters.

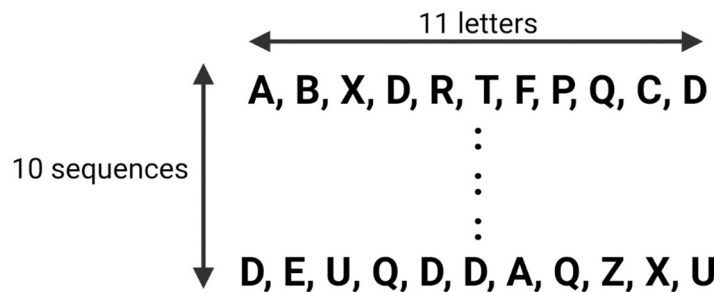


FIGURE 1

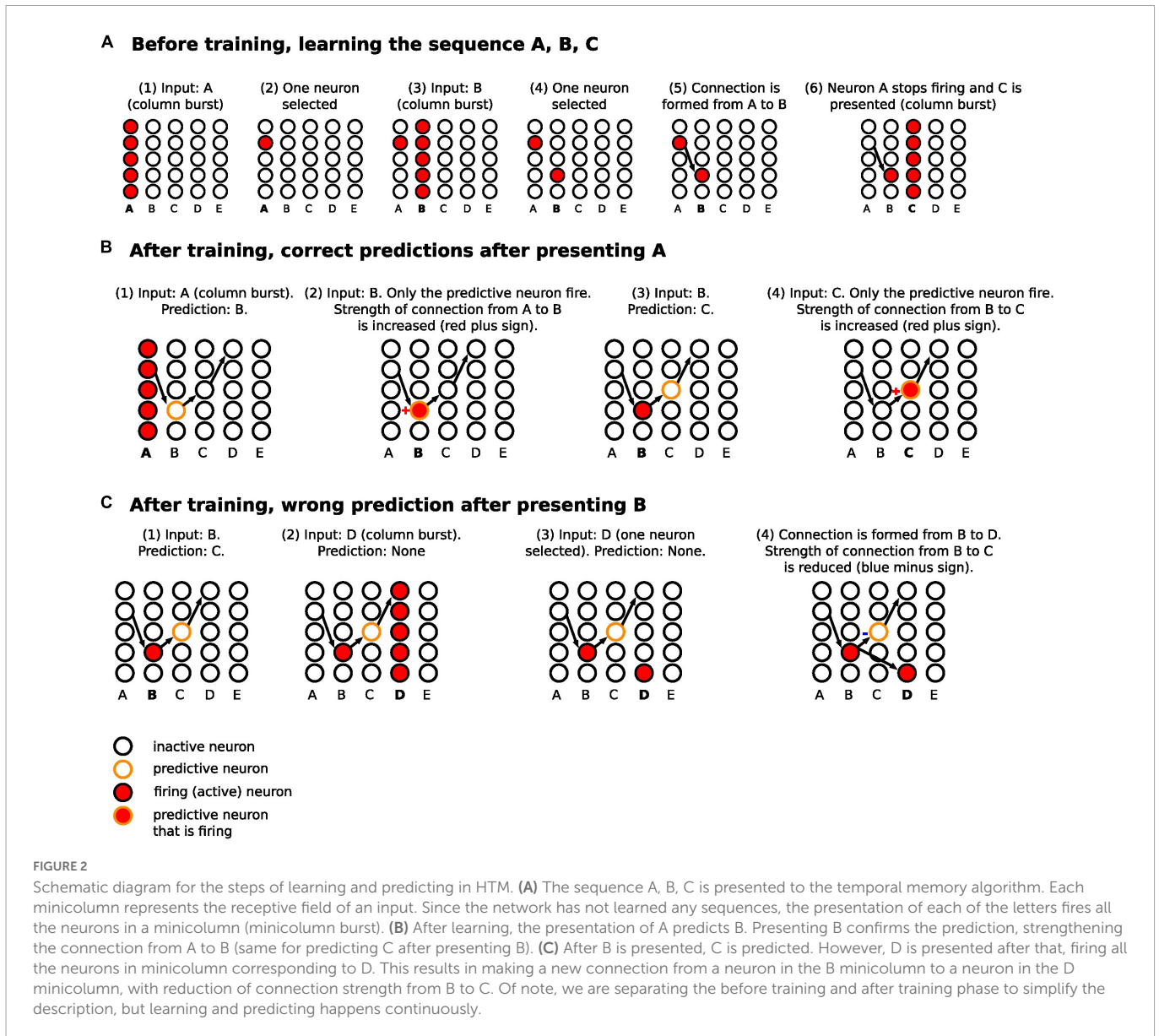
Schematic of the neocortical layer temporal memory (TM) model and the input it learned. The network consisted of 520 columns, with 32 cells per column. Each upper-case letter has a receptive field of 20 columns. The input consists of 10 sequences, each of 11 upper-case letters selected randomly from a uniform distribution with replacement.

functionally connected to its dendritic segment. Spine permanence is a float value that varies from 0 to 1. We used the default value of 0.5 in the simulations. To investigate the effects of varying connected permanence on the networks' performance, we also ran simulations with connected permanence values of 0.25, 0.75, and 0.9. We also examined varying the activation threshold, which is an integer denoting the number of spines needed for dendritic activation. We used the default value of 13 spines in the simulations. And again, to examine the effects of varying activation threshold, we ran additional simulations with activation thresholds of 18 and 20. Since 20 was the maximum number of spines in a fully trained network, running simulations with values higher than 20 would result in no neuronal activation.

Results

The network learned and predicted accurately with training

We presented five input arrays to each of five temporal memory (TM) versions. Each input array consisted of 10 rows \times 11 minicolumns of randomly selected upper-case letters (A-Z) from a uniform distribution. It took each TM network an average of 1,100 iterations to learn to predict the input patterns with 100% prediction accuracy (Figure 3A). For clarity, below we describe separately how the learning took place before and after training. However, the TM learns and predicts continuously.



Before training

The basic learning steps have been described in multiple papers (57, 61, 62) and are illustrated schematically in Figure 2A. In summary, the network learned as follows:

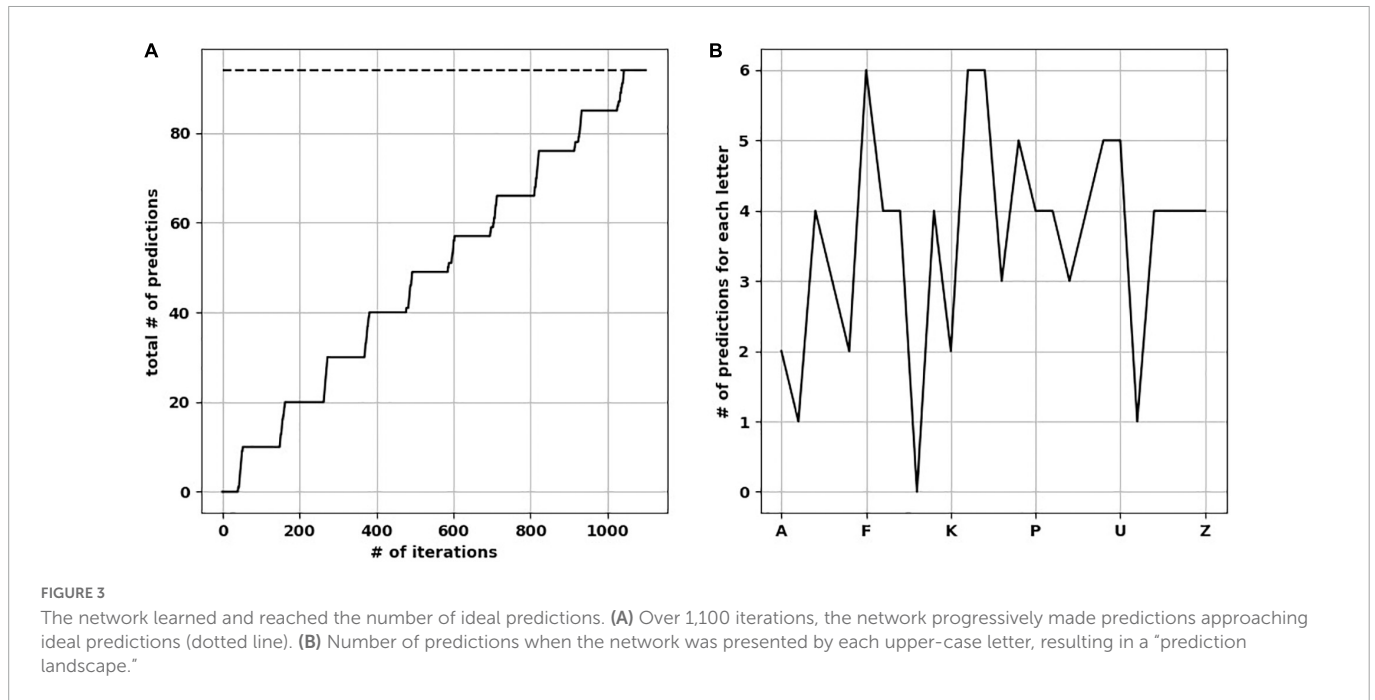
- (1) Before any learning, no predictions were made as the neocortical layer was learning the sequences for the first time.
- (2) When a stimulus was presented (e.g., “A”), all neurons in the minicolumns corresponding to the presented letter, i.e., its receptive fields, were activated.
- (3) Winner neurons in the first set of minicolumns were randomly selected (one neuron from each column).
- (4) With the presentation of the following sequence (e.g., “B”), all neurons in the receptive field minicolumns were activated as well.
- (5) Winner neurons in the second set of minicolumns were randomly selected.

- (6) A connection was made from the winner neurons of the previous time step to the winner neurons in the current time step.
- (7) Repeat from step number two until the sequence ended.

After training

After a network learned the sequences of letters, the network continued predicting and updating its predictions as follows (Figure 2B):

- (1) After the first stimulus was presented, since it was the first one, there were no predictions before it. All neurons in the receptive field minicolumns fired.
- (2) The firing of the neurons across the receptive field set their target neurons into a predictive state. Biologically, this occurs through activation of an NMDA spike that depolarizes the soma



membrane to sub-threshold voltage (60, 63). Through learning, the neurons in the predictive state were in minicolumns corresponding to the next predicted letter.

- (3) When the subsequent input arrived, the next step varied depending on whether the input matched the predicted input, as follows.
 - (a) If the input matched the minicolumns that contained the predictive neurons, then within each minicolumn, the predictive neurons fired earlier than the other neurons in the minicolumn because they were closer to the spiking threshold. This firing in turn inhibited the other neurons in the minicolumn (winner-takes-all), and the predictive neurons were selected as the winner neurons. The synapses between the neurons that were active in the previous time step and the predictive-turned-winner neurons will be strengthened in a Hebbian manner, where neurons that fire repeatedly in close temporal proximity to each other have stronger connections (64).
 - (b) If the input did not match the minicolumns that contained the predictive neurons, all neurons in each minicolumn corresponding to the current input were activated and the winner neurons were randomly selected. New synapses were then formed between the neurons that were active in the previous time step and the randomly selected winner neurons in the current time step. Also, the synapses that were projecting to the incorrectly predicted neurons were weakened. The prediction error signal in these cases was the activation of all neurons in the minicolumns receiving the input.

In the TM model, formation of synaptic connections and changes in their strength is not done by changing the synaptic weights (magnitude of synaptic influence on the neuron). Instead, it is done by modifying synaptic permanence. Permanence is how resistant the

synapses are for removal, where stronger synapses are more resistant to atrophy (57).

The total number of predictions a TM can make

To calculate the total number of predictions each trained TM version is capable of making (with all synapses intact), we presented each upper-case alphabetical letter to the network while pausing new learning to avoid changes in synaptic plasticity. We then calculated how many letters were predicted on the next time step, to create a “prediction landscape” (Figure 3B). The predicted letters were based on the minicolumns containing neurons in the predictive state, and what letter(s) did these minicolumns map to. Since the presented letter and predicted letter(s) form a two-letter sequence (a letter dyad), we compared each prediction landscape to the number of letter dyads (the ground-truth of what the network should be able to predict). As expected, we found no difference in the number of letter dyads between a fully trained network and the frequency of letter dyads in the presented sequences. Counting the predictions across the upper-case letters resulted in the total number of predictions each trained TM makes. The mean number of total predictions across the 25 TM versions was 94.6 (std. dev. = 1.85).

Loss of predictions and increased prediction error with synaptic destruction

To understand how synaptic atrophy can affect the predictive function of a neocortical layer, we randomly destroyed the number of synapses in fractions (25, 50, 60, 70, and 80%) across all trained TM versions (five TM versions trained across five variations of input sequences) and examined the number of predictions made by the models. Reducing the number of synapses resulted in a progressively flatter prediction landscape, starting at the point of

50% synapse destruction and reaching a flat line of zero predictions when there were 20% remaining synapses (Figure 4A). To ensure the reproducibility of this result, we used five random seeds to randomly select the synapses to be destroyed for each fraction of synaptic removal, resulting in 125 network versions for each fraction removed. With 75% of synapses intact, the mean number of predictions was the same as it was at baseline when all synapses were present (Figure 4B, black data points and left y-axis). At 50% removal, the mean number of predictions decreased slightly to 89.1 (std. dev. = 2.7), which is still close to the number of predictions a “healthy” network (100% synapses) had. With 40% of synapses intact, the mean number of predictions dropped to 34.9 (std. dev. = 2.97). With 30% of synapses remaining, the average number of predictions was 2.5 (std. dev. = 1.7), and when only 20% of synapses were remaining, no predictions were generated across the 125 networks.

The reduction we found in the number of predictions with greater degrees of synaptic loss was not accounted for by less activity in the network. On the contrary, with increasing synaptic destruction, more cells were firing as more minicolumns became fully active. We calculated the average confidence the network had with the progressive destruction of synapses by calculating a scaled ratio of winner cells to active cells (Figure 4B, red data points and right y-axis). As more synapses were lost, fewer neurons received the predictions from the active neurons on the previous time step, resulting in more bursts of the minicolumns, denoting prediction error. Thus, the less confident the network is, the more bursting and the more active it is. Interestingly, we found that network confidence dropped much faster than its total number of predictions. With 100% of synapses present, the mean confidence was 0.9 (std. dev. = 0.1), while with 75% of synapses remaining, the mean confidence dropped to 0.68 (std. dev. = 0.06) despite the same number of predictions made. With 50% reduction of synapses, despite a mild reduction in the number of total predictions, the confidence dropped to 0.05 (std. dev. = 0.003).

Loss of confidence but intact number of predictions with dendritic segment destruction

To investigate another possible scenario of connectivity loss, we modeled the destruction of connections along dendritic segments. We found that even after destroying 80% of dendritic segments, the total number of predictions was very similar to the healthy network, unlike with destroying 80% of synapses throughout the network (Figure 5). However, the confidence of the network about the predictions it made dropped markedly, similar to the networks with loss of spines.

Varying hyperparameters

To assess how the findings were dependent on the default hyperparameter values we used, we ran additional simulations while varying two hyperparameters. The first is *connected permanence*, which is the spine permanence above which a spine is connected.

In HTM, synaptic strength is not measured by weight. Instead, permanence is used, i.e., how resistant a synapse is for removal following synaptic depression. Synapses with high permanence would be equivalent to mature mushroom spines. The second hyperparameter we varied was *activation threshold*, which is how many synapses need to be activated on a dendritic segment to trigger subthreshold depolarization, biologically equivalent to generating a dendritic NMDA spike, putting the soma in a predictive state. We found that varying connected permanence did not change the number of predictions or the confidence of either in the intact networks or the “depressed” ones (Figure 6A). Increasing the activation threshold more than the default value of 13 spines did not affect the prediction number or confidence in the intact network. But both the number of predictions and the prediction confidence were reduced as synapses were destroyed (Figure 6B). This was because, with synaptic destruction, there were not enough spines remaining to activate the dendrites, reducing the number of neurons in the predictive state, resulting in lower number of predictions. Also, because there were less neurons in the predictive state, there were more columnar bursts, reducing confidence.

Discussion

The main goal of this study was to mechanistically connect the reduction of synapses seen in postmortem biopsies from MDD patients and in animal models of depression with MDD symptomatology based on the Bayesian brain, predictive coding, and interoception frameworks. Using the temporal memory (TM) model, we ran computer simulations investigating how the prediction making capacity of a neocortical layer would be affected by reducing the number of available synapses since synaptic loss characterizes brains with MDD and restored number of synapses is associated with antidepressant effects.

Emotions are hypothesized to exist on two dimensions (37). The first is valence, which reflects a prediction about net energy balance. A prediction of negative net energy balance, either because of predicted reduced gain or predicted high energy cost, will be perceived as sadness. The second dimension is arousal, which reflects increased prediction error, and is a signal for learning and updating the internal model. MDD was conceptualized as a disorder where interoceptive neocortical regions are stuck in fewer affective states with high prediction errors (37, 45, 65). We started from the first principles of individual neuronal functioning formalized into a functional microcircuit of a single neocortical layer in the HTM framework (57), and we found that reducing the number of synapses would result in *both* a limited number of predictions and increased prediction error.

Reducing the synapses by 50% slightly reduced the number of predictions made, but the confidence strikingly dropped. Reducing the synapses by more than 50% decreased the number of predictions made by the trained networks further in a non-linear fashion. The decrease in prediction numbers was not due to reduced activity following reduced synaptic numbers. On the contrary, neuronal firing increased secondary to increased minicolumn bursts, reflecting higher prediction error (66) (or lower confidence in predictions).

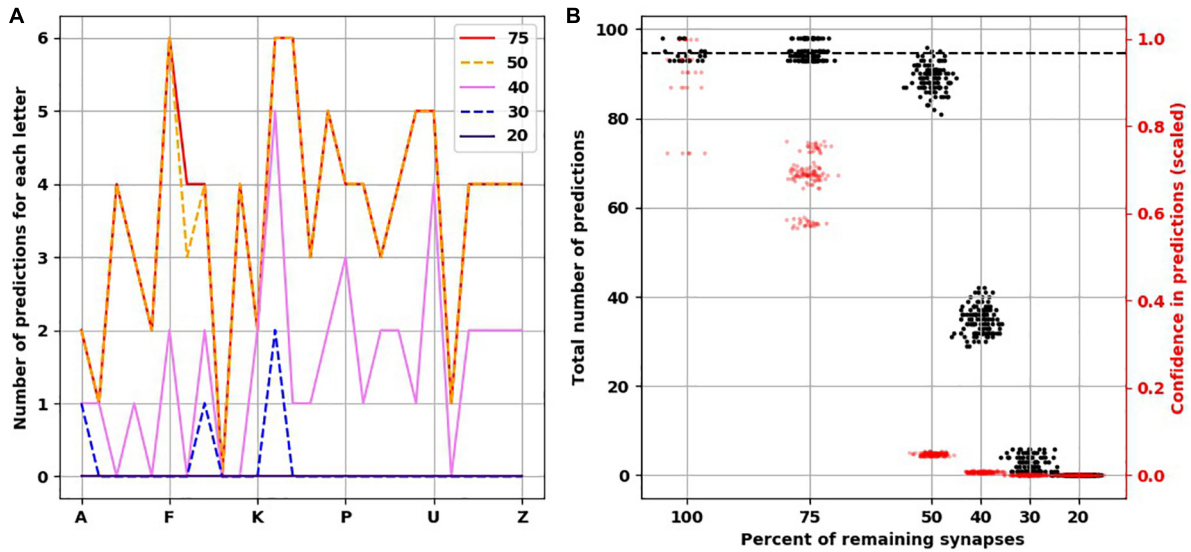


FIGURE 4
 Reduced confidence in predictions before number of predictions is decreased with destroying synapses. **(A)** Illustrative example from the same network in **Figure 3** showing the progressive reduction in number of predictions for each letter with loss of synapses. Each colored line represents the number of remaining synapses. Of note, the number of predictions from network with remaining 75% is identical to network with 100% of synapses. **(B)** At 75% of synapses, the networks encoded similar number of predictions as 100% of synapses, but with less confidence. At 50% of synapses, there is a slight reduction of number of predictions, but the confidence is less than 0.1. There is marked reduction of number of predictions at 40 and 30% of synapses. No predictions could be generated with 20% of synapses. There are 25 simulations at 100%, and 125 simulations at the remaining percentages.

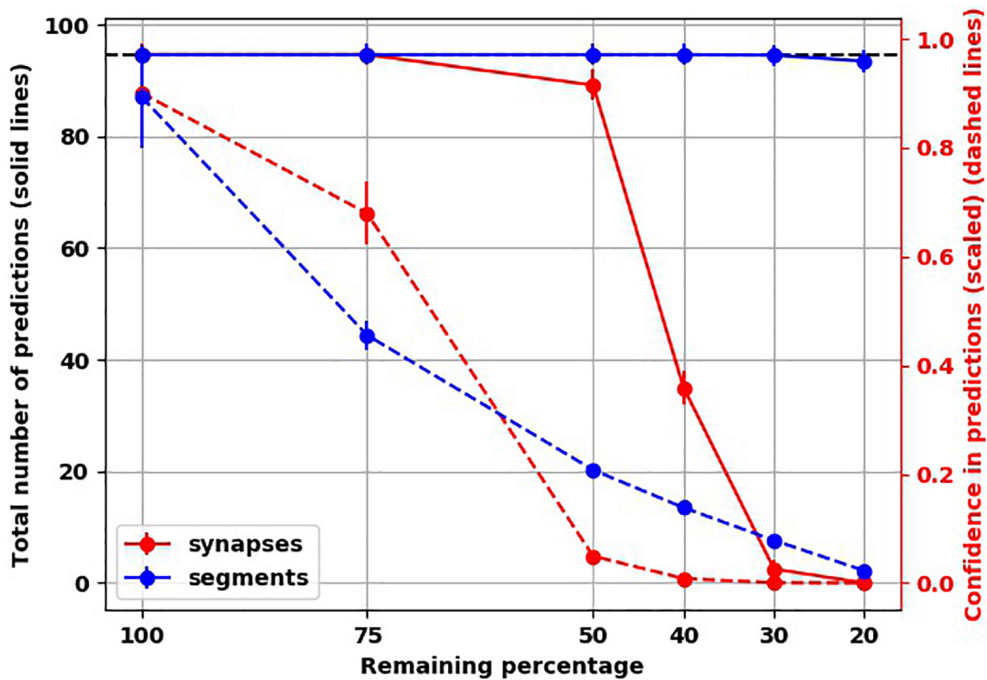
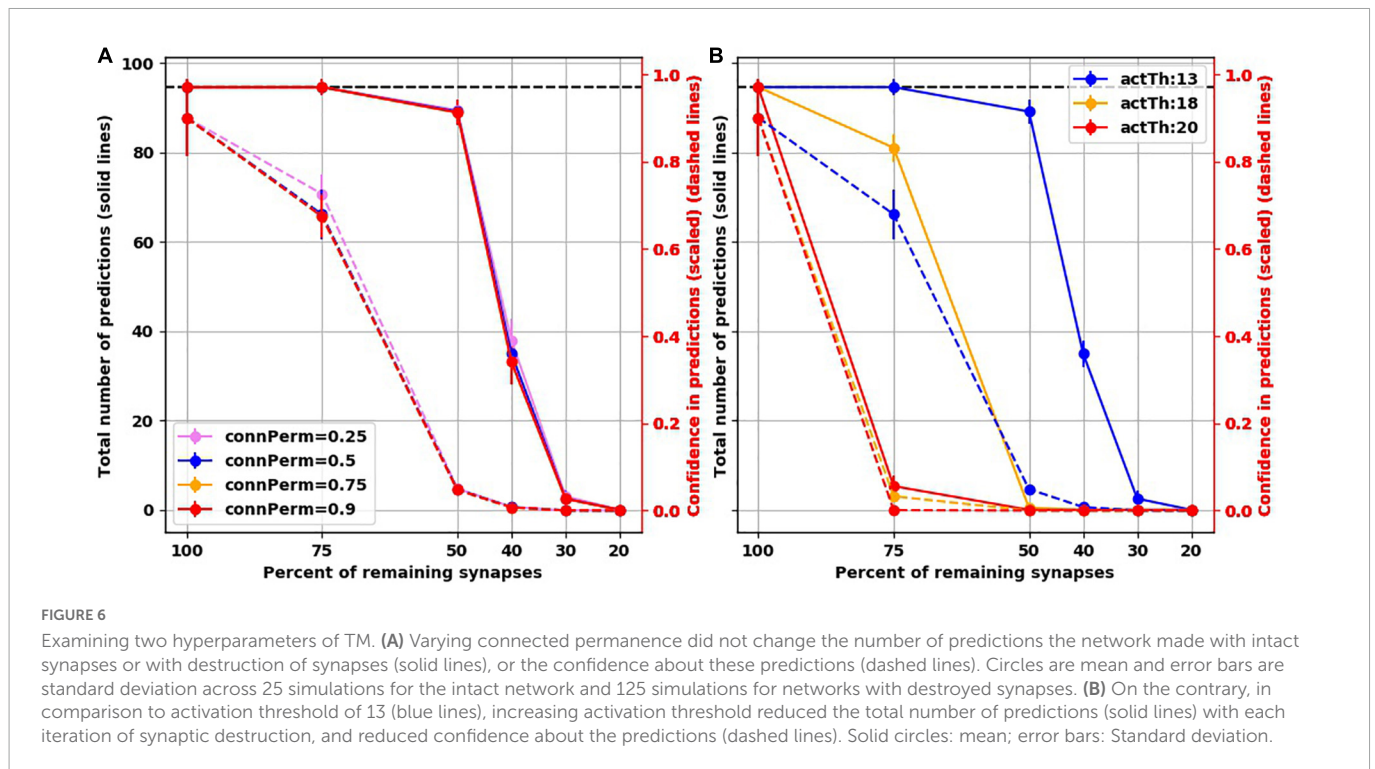


FIGURE 5
 Unlike loss of synapses, loss of connections along dendritic segments affected confidence (dashed lines) but not number of predictions (solid lines).

Clinical relevance and proposed mechanisms

Rumination, where the brain keeps coming back to the same thought content (67), is construed here as a manifestation of limited predictions and reduced confidence in these predictions.

Being stuck in a negative mood could be considered “affective rumination,” or more formally, affective inertia (68). Koval et al. (67) found that the symptom of being cognitively stuck (i.e., rumination) and emotionally stuck (emotional inertia) are positively correlated with one another yet independent in determining the overall severity of the clinical depression. According to the



Embodied Predictive Interoception Coding (EPIC) model, the brain’s interoceptive agranular cortices (44) project diffusely to other cortical areas to relay affective states (69, 70). Applying our results to the agranular cortices, e.g., subgenual cingulate cortex (sgACC), suggests a mechanism by which a condition like depression with reduced synapses has limited interoceptive predictive states and high prediction errors, reflecting limited mood experiences, i.e., constricted/restricted affect but with high arousal and perhaps heightened anxiety levels (71, 72), shaping further expectations of patients (73).

Our study also suggests why the limited predicted states in MDD are experienced as negative affective states. According to the concept of interoceptive allostasis, states with predictions of negative energy balance would be perceived as affectively unfavorable (45). When there is reduced confidence in predictions made by interoceptive cortices, the uncertainty would be transmitted to other neocortical and subcortical regions (74, 75). Such uncertainty might signal the need to be “ready for anything,” to prepare for a larger number of possibilities. However, being ready for anything would also predict high energy costs and negative energy balance, which would be perceived by the depressed patient as negative affect.

The non-linear relationship of our findings could explain the rapid antidepressant properties of ketamine and psilocybin. Above a threshold, restoring a fraction of synapses would be enough to improve the confidence in the predictions made, reducing the prediction error signal, thus limiting the number of possibilities for which energy costs have to be calculated, and ultimately reducing the negative energy balance.

Our findings also suggest a possible mechanism for how atrophy in cortical regions in MDD (76–79) could be associated with increased neuronal activity. Multiple studies showed a spontaneous state of hypermetabolism of sgACC in depression, particularly in patients struggling with treatment-refractory depression; this biochemical marker reflecting increased neuronal activity has been

observed with positron-emission tomography (PET) (80) and in human MEG studies (81), and is similar to observations arising in mouse models of depression (82). Our results shows that synaptic and dendritic atrophy would lead to less neurons in predictive states, resulting in more columnar bursting, and thus increased neuronal firing.

This work illustrates how to formalize Bayesian brain theories into a neocortical neurobiological framework. For example, suicidal ideations can result from a limited ability to consider alternative future scenarios to the extent that the patient cannot escape internal pain, a construct called internal entrapment. De Beurs et al. (83) investigated the interactions of two prominent theories underlying suicidality, namely interpersonal psychological theory (IPT) and the integrated motivational-volitional (IMV) model. They found that two factors, one from each theory, were directly related to suicidal ideations; the factors were *perceived burdensomeness* of IPT and *internal entrapment* from the IMV framework. Subjective experiences are predictions based on internal and external cues. To be adaptive, subjective experiences need to be updated when the cues change. In light of our findings, failure to update subjective states, such as perceived burdensomeness, would be related to the restricted ability to make new predictions secondary to the limited number of synapses. Internal entrapment describes the inability to escape internal pain, which could be conceptualized as a limited view of new possibilities because of “tunnel vision” (84). Such a limited view would be expected when motivation visceromotor cortices are unable to make more predictions because of a reduced number of spines and/or decreased spine maturation. These spine changes are supported by studies of brain derived neurotrophic factor (BDNF), a growth factor implicated in plasticity and spine growth, in postmortem brains of people who died by suicide (85, 86). In fact, the notion that functional synaptic connection may be the final common pathophysiologic mediator of depression (87) is consistent with our synapse model presented here, and is also suggested by the following:

- (1) Stress-based animal models induce BDNF methylation resulting in decreased expression, while antidepressants renew BDNF production through acetylation (88).
- (2) Likewise, in humans, decreased BDNF serum levels are seen in depressed patients, and these levels correct to the level of healthy controls with selective-serotonin reuptake inhibitors (SSRIs) (89).
- (3) Synaptic density is decreased in patients with depression, detected by positron emission tomography (PET) (90), while BDNF induces dendritic arborization (91) and neurogenesis, which underlies reversal of depressive behavior in animal models after 1-month of SSRIs (92) or tricyclic antidepressants (93).
- (4) Electroconvulsive therapy also enhances BDNF production, spine maturation, and neurogenesis (13, 14).

Therefore, synapses and BDNF are reduced in depression, but restoration of BDNF rescues synaptogenesis, reversing depression.

Relation to other models

Other computer modeling work of MDD has shown how cortical regions can be caught in their dynamic states in MDD. A computer model was developed of two interacting regions, ventral anterior cingulate cortex (vACC) engaged in emotional processing, and DLPFC involved in cognitive processing. Ramirez-Mahaluf et al. used the model to investigate how the two regions interacted together in depression (94). Each region contained 800 pyramidal neurons and 200 interneurons, reciprocally connected. The two regions inhibited each other by activating the interneuronal population in the other region. They modeled a healthy condition where each region can either be active in a state of low-rate asynchronous firing (0.5–1 spikes/second) or high-rate asynchronous firing (25–30 spikes/second). The corresponding region became active in the healthy network when presented with either a cognitive (working memory) or an emotional (sadness provocation) task. When they modeled depression as slower glutamate reuptake in vACC, it resulted in sustained activity of vACC that DLPFC could not inhibit. Thus, the network could not switch to active DLPFC when presented with the cognitive task. They explored the progression of MDD symptom severity from mild to moderate to severe, and found that switching to active DLPFC functioning in the face of a cognitive task was more difficult as the disease progressed. They examined the effect of SSRIs on the dynamics of switching and found that SSRIs inhibited the excitatory neuronal population in the vACC and reversed the dynamics back to the healthy bistable state, though the reversal became more difficult in more severe depression. When they modeled the effects of deep brain stimulation (DBS) on their model, they found that DBS also reversed the dynamics. Similar to our study, modeling depression resulted in increased activity of the interoceptive region model (vACC), where the increased firing was secondary to increased prediction errors.

Somatostatin-positive interneurons (SST) are a population of interneurons that mainly target the distal apical dendrites of pyramidal neurons in the neocortex. Deficiency of SST interneurons might play a role in MDD (95). To examine the effects of reducing somatostatin (SST) mRNA in postmortem tissue from depressed patients (96), Yao et al. (97) developed a computer model of neocortical layer 2/3 based on both human and rodent data. The model consisted of pyramidal neurons targeting three interneuronal populations: SST, parvalbumin-positive (PV),

and vasoactive intestinal peptide-positive (VIP) interneurons. SST interneurons targeted the apical dendrites of pyramidal neurons, as well as the PV and VIP interneurons. To model depression, they decreased the SST inhibition by 40% and examined how that affected neuronal firing and stimulus processing. SST activity reduction increased the background firing rates of the different neuronal populations, including the pyramidal neurons. However, when they presented a stimulus to their model, the stimulus triggered similar pyramidal neuronal firing rates. Similar to our findings, the signal-to-noise ratio in that model (calculated as the ratio between baseline pyramidal firing rate and the post-stimulus firing rate) was reduced because the baseline firing rate was higher in the depressed network model than in the healthy network model. They also found that reduction in the signal-to-noise ratio doubled the rates of both failed detection and false detection of the stimulus when it was presented in the depressed models.

In the light of evolutionary psychology and the Bayesian brain framework, one theory conceptualizes depression as a way to minimize surprises in social interactions (98). This theory is based on observations that patients diagnosed with depression have learned to expect negative interactions with others based on their past experiences. This would reduce surprises (i.e., prediction error), perhaps as a mechanism to reduce emotional vulnerability when such negative interactions arise. At the same time, this theory assumes that it is challenging for a depressed patient to update their (negatively valenced) beliefs. Given their predictions about negative interactions, it is possible that depressed patients tend to avoid social interactions, and thus miss interactions that could update their expectations about the outcome of the social interaction. Another possibility could be that neuromodulatory neurobiological factors in a depressed brain, such as reward learning mediated by dopamine neurotransmission, make it difficult for patients to update their predictions.

The findings from our study hint at a third potential mechanism underlying the difficulty a depressed patient may have in updating their negative cognitions. Our results show that a reduced number of synapses to encode more predictions will limit the ability to update one's beliefs with novel interactions or experiences. Typically, authors of prior studies with Bayesian brain formulations of psychiatric disorders [e.g., (98, 99)] have assumed that a distinct neuronal population be used for calculating predictions than the population used for calculating prediction errors. Barrett and Simmons (44) suggested that agranular layers of interoceptive cortices can send predictions but lack the number of neurons needed to compute prediction errors of interoceptive inputs. In the TM model of the HTM framework that we used here, neurons within a single neocortical layer were capable of both making predictions and calculating prediction errors. This suggests that agranular cortical areas can both send predictions and compute prediction errors.

Future directions

This is a proof-of-concept study of a single neocortical layer exhibiting synaptic atrophy as seen in MDD and animal models of depression. David Marr suggested three levels of analysis when looking at the function of neuronal microcircuits (100, 101). The first is a computational level: What is the problem that the microcircuit solves? The second level is the algorithmic level, describing the steps needed to solve the problem. The third level is the implementation level, how the microcircuit elements, such as neuronal populations

and their connections, perform the steps of the algorithm. We aimed to connect Marr's three levels of analysis in the context of MDD using the HTM framework. At the computational level, the problem is making predictions about energy expenditures in a changing environment. At the algorithmic level, the network represents multiple possible predictions and selects the ones which agree with the current input (57). At the implementation level, the neurons used are restrained by the neurobiology of neocortical neurons, e.g., sub-threshold soma depolarization by NMDA spikes when clustered synapses are activated within a short time window (60, 102).

To validate the main prediction from this work, we can look at sequences of action potentials, or replays. According to the HTM framework, predictions are encoded as temporal firing sequences. Replay of neocortical temporal sequences has been shown widely across multiple neocortical regions, usually in the context of memory retrieval (103, 104), as well as in neocortical slices (105). Replay could be examined in animal models of depression, and from invasive recordings in patients undergoing DBS for depression, as in Scangos et al. (106). Our finding of reduced number of predictions would mean that there are less potential sequences of neuronal activation in a network, reflected in depression as reduced replay variability.

To examine other aspects of depression pathophysiology, there are several ways to take this work further. The dendrites represented here are basal. Within the structural model of neocortical organization, agranular cortical areas receive input from dysgranular cortical areas (with less developed layer IV) onto deeper layers (107), most likely terminating on basal dendrites of layer 5 and 6 neurons. Both ketamine (108) and psilocybin (36) increased the number of spines along the basal dendrites, as well as along the apical dendrites [e.g., (31, 35)]. Adding apical dendrites to the model will allow us to examine how apical synaptic atrophy affects the predictive function of the neocortex, especially with the postulated role of SST interneurons targeting apical dendrites in depression (96).

Laminar structure has been suggested to play a role in the perception-action cycle through integration of information across layers within cortical columns (109, 110). Expanding the model to include multiple HTM layers will allow us to investigate the effects of synaptic loss within layers vs. between layers on predictions made by a whole cortical column (111) and to examine how interlaminar information flow changes in depression. Modeling more than one region, including subcortical nuclei, would also enable us to predict how manipulating neurons and microcircuits, usually through pharmacology, would affect elements of Bayesian inference underlying symptoms, thus creating a mechanistic connection between therapeutic interventions and symptom improvement. Yet another future direction would be to implement neuromodulators like monoamines that play a role in depression. Our model does not take the effects of endogenous neuromodulators such as serotonin or dopamine into account. The level of modeling we did here also does not consider the immunological and hormonal mechanisms that mediate the effects of chronic stress on synaptic atrophy. Instead, we looked at one of the possible outcomes of chronic stress, namely synaptic atrophy, and how this could be related to MDD.

To our knowledge, this is the first study applying algorithms from the biologically constrained HTM framework to neocortical abnormalities found in a psychiatric disorder. This work highlights the potential of the HTM framework to formalize the predictive

capacity of neocortical layers and the possible insights this line of investigation can bring to understanding psychiatric disorders. Instead of limiting brain microcircuit modeling to molecular and microcircuit mechanisms underlying specific neural dynamics, this approach extends microcircuit modeling into the functional realm, potentially connecting microcircuit models to behavioral tasks and symptoms (112–114), using neurons that have more biological features than ones traditionally used in artificial neural networks.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

MS, MK, and RS contributed to the conception of the study. MS and JB contributed to the design of the study. MS ran the simulations, did the data analysis, and wrote the first draft of the manuscript. MK, RS, JB, and LC wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

MS was supported by the National Institute of Mental Health (NIMH) (R21MH125199) and by the National Institute of General Medical Sciences (NIGMS) (P20GM103645). JB was supported by the National Institute of General Medical Sciences (NIGMS) (P20GM130452).

Conflict of interest

MS was a consultant for In Silico Biosciences, Inc. LC has consulted for Otsuka, Sunovion, Neuronetics, Nexstim, Janssen, Affect Neuro, Sage Therapeutics, and Neuroief and also received research support through clinical trial contracts at Butler Hospital with Janssen, Neuronetics, Affect Neuro, and Neuroief.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Whiteford H, Degenhardt L, Rehm J, Baxter A, Ferrari A, Erskine H, et al. Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. *Lancet*. (2013) 382:1575–86. doi: 10.1016/S0140-6736(13)61611-6
- Kessler R, Berglund P, Demler O, Jin R, Koretz D, Merikangas K, et al. The epidemiology of major depressive disorder: results from the national comorbidity survey replication (NCS-R). *JAMA*. (2003) 289:3095–105. doi: 10.1001/jama.289.23.3095
- Kessler RC. The costs of depression. *Psychiatr Clin North Am*. (2012) 35:1–14. doi: 10.1016/j.psc.2011.11.005
- Rosoklija G, Toomayan G, Ellis SP, Keilp J, Mann JJ, Latov N, et al. Structural abnormalities of subicular dendrites in subjects with schizophrenia and mood disorders: preliminary findings. *Arch Gen Psychiatry*. (2000) 57:349–56. doi: 10.1001/archpsyc.57.4.349
- Kang H, Voleti B, Hajszan T, Rajkowska G, Stockmeier C, Licznarski P, et al. Decreased expression of synapse-related genes and loss of synapses in major depressive disorder. *Nat Med*. (2012) 18:1413–7. doi: 10.1038/nm.2886
- Kuhn M, Mainberger F, Feige B, Maier J, Wirminghaus M, Limbach L, et al. State-dependent partial occlusion of cortical LTP-like plasticity in major depression. *Neuropsychopharmacology*. (2016) 41:1521–9. doi: 10.1038/npp.2015.310
- Cook SC, Wellman CL. Chronic stress alters dendritic morphology in rat medial prefrontal cortex. *J Neurobiol*. (2004) 60:236–48. doi: 10.1002/neu.20025
- Radley JJ, Rocher AB, Miller M, Janssen WG, Liston C, Hof PR, et al. Repeated stress induces dendritic spine loss in the rat medial prefrontal cortex. *Cereb Cortex*. (2006) 16:313–20. doi: 10.1093/cercor/bhi104
- Radley JJ, Sisti HM, Hao J, Rocher AB, McCall T, Hof PR, et al. Chronic behavioral stress induces apical dendritic reorganization in pyramidal neurons of the medial prefrontal cortex. *Neuroscience*. (2004) 125:1–6. doi: 10.1016/j.neuroscience.2004.01.006
- Bessa JM, Ferreira D, Melo I, Marques F, Cerqueira JJ, Palha JA, et al. The mood-improving actions of antidepressants do not depend on neurogenesis but are associated with neuronal remodeling. *Mol Psychiatry*. (2009) 14:764–73. doi: 10.1038/mp.2008.119
- Vlachos A, Müller-Dahlhaus F, Roskopp J, Lenz M, Ziemann U, Deller T. Repetitive magnetic stimulation induces functional and structural plasticity of excitatory postsynapses in mouse organotypic hippocampal slice cultures. *J Neurosci*. (2012) 32:17514–23. doi: 10.1523/JNEUROSCI.0409-12.2012
- Cambiaghi M, Cherchi L, Masin L, Infortuna C, Briski N, Caviasso C, et al. High-frequency repetitive transcranial magnetic stimulation enhances layer II/III morphological dendritic plasticity in mouse primary motor cortex. *Behav Brain Res*. (2021) 410:113352. doi: 10.1016/j.bbr.2021.113352
- Maynard KR, Hobbs JW, Rajpurohit SK, Martinowich K. Electroconvulsive seizures influence dendritic spine morphology and BDNF expression in a neuroendocrine model of depression. *Brain Stimul*. (2018) 11:856–9. doi: 10.1016/j.brs.2018.04.003
- Zhao C, Warner-Schmidt J, Duman RS, Gage FH. Electroconvulsive seizure promotes spine maturation in newborn dentate granule cells in adult rat. *Dev Neurobiol*. (2012) 72:937–42. doi: 10.1002/dneu.20986
- Brown JC, DeVries WH, Korte JE, Sahlem GL, Bonilha L, Short EB, et al. NMDA receptor partial agonist, d-cycloserine, enhances 10 Hz rTMS-induced motor plasticity, suggesting long-term potentiation (LTP) as underlying mechanism. *Brain Stimul*. (2020) 13:530–2. doi: 10.1016/j.brs.2020.01.005
- Brown JC, Yuan S, DeVries WH, Armstrong NM, Korte JE, Sahlem GL, et al. NMDA-receptor agonist reveals LTP-like properties of 10-Hz rTMS in the human motor cortex. *Brain Stimul*. (2021) 14:619–21. doi: 10.1016/j.brs.2021.03.016
- George MS, Lisanby SH, Avery D, McDonald WM, Durkalski V, Pavlicova M, et al. Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: a sham-controlled randomized trial. *Arch Gen Psychiatry*. (2010) 67:507–16. doi: 10.1001/archgenpsychiatry.2010.46
- Carpenter LL, Janicak PG, Aaronson ST, Boyadjis T, Brock DG, Cook AI, et al. Transcranial magnetic stimulation (TMS) for major depression: a multisite, naturalistic, observational study of acute treatment outcomes in clinical practice. *Depress Anxiety*. (2012) 29:587–96. doi: 10.1002/da.21969
- O'Reardon JP, Solvason HB, Janicak PG, Sampson S, Isenberg KE, Nahas Z, et al. Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. *Biol Psychiatry*. (2007) 62:1208–16. doi: 10.1016/j.biopsych.2007.01.018
- Cole J, Selby B, Ismail Z, McGirr A. D-cycloserine normalizes long-term motor plasticity after transcranial magnetic intermittent theta-burst stimulation in major depressive disorder. *Clin Neurophysiol*. (2021) 132:1770–6. doi: 10.1016/j.clinph.2021.04.002
- Cole J, Sohn MN, Harris AD, Bray SL, Patten SB, McGirr A. Efficacy of adjunctive d-cycloserine to intermittent theta-burst stimulation for major depressive disorder: a randomized clinical trial. *JAMA Psychiatry*. (2022) 79:1153–61. doi: 10.1001/jamapsychiatry.2022.3255
- Berman RM, Cappiello A, Anand A, Oren DA, Heninger GR, Charney DS, et al. Antidepressant effects of ketamine in depressed patients. *Biol Psychiatry*. (2000) 47:351–4. doi: 10.1016/S0006-3223(99)00230-9
- Zarate CA, Singh JB, Carlson PJ, Brutsche NE, Ameli R, Luckenbaugh DA, et al. A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch Gen Psychiatry*. (2006) 63:856–64. doi: 10.1001/archpsyc.63.8.856
- Murrough J, Perez A, Pillemer S, Stern J, Parides M, aan het Rot M, et al. Rapid and longer-term antidepressant effects of repeated ketamine infusions in treatment-resistant major depression. *Biol Psychiatry*. (2013) 74:250–6. doi: 10.1016/j.biopsych.2012.06.022
- Singh J, Fedgchin M, Daly E, De Boer P, Cooper K, Lim P, et al. A double-blind, randomized, placebo-controlled, dose-frequency study of intravenous ketamine in patients with treatment-resistant depression. *Am J Psychiatry*. (2016) 173:816–26. doi: 10.1176/appi.ajp.2016.16010037
- Singh J, Fedgchin M, Daly E, Xi L, Melman C, De Bruecker G, et al. Intravenous esketamine in adult treatment-resistant depression: a double-blind, double-randomization, placebo-controlled study. *Biol Psychiatry*. (2016) 80:424–31. doi: 10.1016/j.biopsych.2015.10.018
- Ross S, Bossis A, Guss J, Agin-Liebes G, Malone T, Cohen B, et al. Rapid and sustained symptom reduction following psilocybin treatment for anxiety and depression in patients with life-threatening cancer: a randomized controlled trial. *J Psychopharmacol*. (2016) 30:1165–80. doi: 10.1177/0269881116675512
- Phoumthippavong V, Barthas F, Hassett S, Kwan A. Longitudinal effects of ketamine on dendritic architecture in vivo in the mouse medial frontal cortex. *eNeuro*. (2016) 3. doi: 10.1523/ENEURO.0133-15.2016
- Li Q, Yan H, Wilson W, Swartzwelder H. Modulation of NMDA and AMPA-mediated synaptic transmission by CB1 receptors in frontal cortical pyramidal cells. *Brain Res*. (2010) 1342:127–37. doi: 10.1016/j.brainres.2010.04.029
- Li N, Liu R, Dwyer J, Banas M, Lee B, Son H, et al. Glutamate N-methyl-D-aspartate receptor antagonists rapidly reverse behavioral and synaptic deficits caused by chronic stress exposure. *Biol Psychiatry*. (2011) 69:754–61. doi: 10.1016/j.biopsych.2010.12.015
- Ng L, Huang Y, Han L, Chang R, Chan Y, Lai CSW. Ketamine and selective activation of parvalbumin interneurons inhibit stress-induced dendritic spine elimination. *Transl Psychiatry*. (2018) 8:272. doi: 10.1038/s41398-018-0321-5
- Ota KT, Liu RJ, Voleti B, Maldonado-Aviles JG, Duric V, Iwata M, et al. REDD1 is essential for stress-induced synaptic loss and depressive behavior. *Nat Med*. (2014) 20:531–5.
- Knott GW, Holtmaat A, Wilbrecht L, Welker E, Svoboda K. Spine growth precedes synapse formation in the adult neocortex in vivo. *Nat Neurosci*. (2006) 9:1117–24. doi: 10.1038/nn1747
- Moda-Sava R, Murdock M, Parekh P, Fetcho R, Huang B, Huynh T, et al. Sustained rescue of prefrontal circuit dysfunction by antidepressant-induced spine formation. *Science*. (2019) 364:eaat8078. doi: 10.1126/science.aat8078
- Fuchikami M, Thomas A, Liu R, Wohleb ES, Land BB, DiLeone RJ, et al. Optogenetic stimulation of infralimbic PFC reproduces ketamine's rapid and sustained antidepressant actions. *Proc Natl Acad Sci USA*. (2015) 112:8106–11. doi: 10.1073/pnas.1414728112
- Shao LX, Liao C, Gregg I, Davoudian PA, Savalia NK, Delagarza K, et al. Psilocybin induces rapid and persistent growth of dendritic spines in frontal cortex in vivo. *Neuron*. (2021) 109:2535–44.e4. doi: 10.1016/j.neuron.2021.06.008
- Barrett LF, Quigley KS, Hamilton P. An active inference theory of allostasis and interoception in depression. *Philos Trans R Soc Lond B Biol Sci*. (2016) 371. doi: 10.1098/rstb.2016.0011
- Petzschner F, Weber L, Gard T, Stephan K. Computational psychosomatics and computational psychiatry: toward a joint framework for differential diagnosis. *Biol Psychiatry*. (2017) 82:421–30. doi: 10.1016/j.biopsych.2017.05.012
- McEwen BS. Mood disorders and allostatic load. *Biol Psychiatry*. (2003) 54:200–7. doi: 10.1016/S0006-3223(03)00177-X
- McEwen BS. Protection and damage from acute and chronic stress: allostasis and allostatic overload and relevance to the pathophysiology of psychiatric disorders. *Ann N Y Acad Sci*. (2004) 1032:1–7.
- Walsh KS, McGovern DP, Clark A, O'Connell RG. Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Ann N Y Acad Sci*. (2020) 1464:242–68. doi: 10.1111/nyas.14321
- Millidge B, Seth A, Buckley CL. Predictive coding: a theoretical and experimental review. *arXiv*. (2021) [Preprint] arXiv: 2107.12979
- Adams RA, Shipp S, Friston KJ. Predictions not commands: active inference in the motor system. *Brain Struct Funct*. (2013) 218:611–43.
- Barrett LF, Simmons WK. Interoceptive predictions in the brain. *Nat Rev Neurosci*. (2015) 16:419–29. doi: 10.1038/nrn3950
- Barrett LF, Satpute AB. Historical pitfalls and new directions in the neuroscience of emotion. *Neurosci Lett*. (2019) 693:9–18. doi: 10.1016/j.neulet.2017.07.045
- Mayberg HS. Limbic-cortical dysregulation: a proposed model of depression. *J Neuropsychiatry Clin Neurosci*. (1997) 9:471–81. doi: 10.1176/jnp.9.3.471

47. Figue M, Mayberg H. The future of personalized brain stimulation. *Nat Med.* (2021) 27:196–7. doi: 10.1038/s41591-021-01243-7
48. Fox MD, Buckner RL, White MP, Greicius MD, Pascual-Leone A. Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate. *Biol Psychiatry.* (2012) 72:595–603.
49. Weigand A, Horn A, Caballero R, Cooke D, Stern AP, Taylor SF, et al. Prospective validation that subgenual connectivity predicts antidepressant efficacy of transcranial magnetic stimulation sites. *Biol Psychiatry.* (2018) 84:28–37. doi: 10.1016/j.biopsych.2017.10.028
50. Cole EJ, Phillips AL, Bentzley BS, Stimpson KH, Nejad R, Barmak F, et al. Stanford neuromodulation therapy (SNT): a double-blind randomized controlled trial. *Am J Psychiatry.* (2021) 179:132–41. doi: 10.1176/appi.ajp.2021.20101429
51. Cash RFH, Zalesky A, Thomson RH, Tian Y, Cocchi L, Fitzgerald PB. Subgenual functional connectivity predicts antidepressant treatment response to transcranial magnetic stimulation: independent validation and evaluation of personalization. *Biol Psychiatry.* (2019) 86:e5–7. doi: 10.1016/j.biopsych.2018.12.002
52. Lashley KS. *The Problem of Serial Order in Behavior.* Oxford: Bobbs-Merrill (1951).
53. Rosenbaum DA, Cohen RG, Jax SA, Weiss DJ, van der Wel R. The problem of serial order in behavior: Lashley's legacy. *Hum Mov Sci.* (2007) 26:525–54. doi: 10.1016/j.humov.2007.04.001
54. Grossberg S. Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world. *Neural Netw.* (2013) 37:1–47. doi: 10.1016/j.neunet.2012.09.017
55. Larkum M. A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci.* (2013) 36:141–51.
56. Mountcastle VB. The columnar organization of the neocortex. *Brain.* (1997) 120(Pt 4):701–22. doi: 10.1093/brain/120.4.701
57. Hawkins J, Ahmad S. Why neurons have thousands of synapses, a theory of sequence memory in neocortex. *Front Neural Circuits.* (2016) 10:23.
58. George D, Hawkins J. Towards a mathematical theory of cortical micro-circuits. *PLoS Comput Biol.* (2009) 5:e1000532. doi: 10.1371/journal.pcbi.1000532
59. Hole K, Ahmad S. A thousand brains: toward biologically constrained AI. *SN Appl Sci.* (2021) 3:1–14. doi: 10.1007/s42452-021-04715-0
60. Antic SD, Zhou WL, Moore AR, Short SM, Ikonomu KD. The decade of the dendritic NMDA spike. *J Neurosci Res.* (2010) 88:2991–3001. doi: 10.1002/jnr.22444
61. Hawkins J, Ahmad S, Cui Y. A theory of how columns in the neocortex enable learning the structure of the world. *Front Neural Circuits.* (2017) 11:81. doi: 10.3389/fncir.2017.00081
62. Hawkins J, Lewis M, Klukas M, Purdy S, Ahmad S. A framework for intelligence and cortical function based on grid cells in the neocortex. *Front Neural Circuits.* (2019) 12:121. doi: 10.3389/fncir.2018.00121
63. Schiller J, Major G, Koester HJ, Schiller Y. NMDA spikes in basal dendrites of cortical pyramidal neurons. *Nature.* (2000) 404:285–9. doi: 10.1038/35005094
64. Hebb DO. The first stage of perception: growth of the assembly. *Org Behav.* (1949) 4:60–78.
65. Holtzheimer PE, Mayberg HS. Stuck in a rut: rethinking depression and its treatment. *Trends Neurosci.* (2011) 34:1–9. doi: 10.1016/j.tins.2010.10.004
66. Byrne F. Encoding reality: prediction-assisted cortical learning algorithm in hierarchical temporal memory. *arXiv.* (2015). [Preprint] arXiv: 1509.08255
67. Koval P, Kuppens P, Allen NB, Sheeber L. Getting stuck in depression: the roles of rumination and emotional inertia. *Cogn Emot.* (2012) 26:1412–27. doi: 10.1080/02699931.2012.667392
68. Suls J, Green P, Hillis S. Emotional reactivity to everyday problems, affective inertia, and neuroticism. *Pers Soc Psychol Bull.* (1998) 24:127–36. doi: 10.1177/0146167298242002
69. Barbas H. General cortical and special prefrontal connections: principles from structure to function. *Annu Rev Neurosci.* (2015) 38:269–89. doi: 10.1146/annurev-neuro-071714-033936
70. Garcia-Cabezas M, Zikopoulos B, Barbas H. The structural model: a theory linking connections, plasticity, pathology, development and evolution of the cerebral cortex. *Brain Struct Funct.* (2019) 224:985–1008.
71. Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med.* (2017) 23:28–38. doi: 10.1038/nm.4246
72. Siddiqi SH, Taylor SF, Cooke D, Pascual-Leone A, George MS, Fox MD. Distinct symptom-specific treatment targets for circuit-based neuromodulation. *Am J Psychiatry.* (2020) 177:435–46. doi: 10.1176/appi.ajp.2019.19090915
73. Eldar E, Rutledge RB, Dolan RJ, Niv Y. Mood as representation of momentum. *Trends Cogn Sci.* (2016) 20:15–24. doi: 10.1016/j.tics.2015.07.010
74. Kumar S, Black SJ, Hultman R, Szabo ST, DeMaio KD, Du J, et al. Cortical control of affective networks. *J Neurosci.* (2013) 33:1116–29. doi: 10.1523/JNEUROSCI.0092-12.2013
75. Kumar S, Hultman R, Hughes D, Michel N, Katz BM, Dzirasa K. Prefrontal cortex reactivity underlies trait vulnerability to chronic social defeat stress. *Nat Commun.* (2014) 5:4537–4537. doi: 10.1038/ncomms5537
76. Stockmeier CA, Rajkowska G. Cellular abnormalities in depression: evidence from postmortem brain tissue. *Dial Clin Neurosci.* (2004) 6:185–97. doi: 10.31887/DCNS.2004.6.2/cstockmeier
77. Dunlop BW, Mayberg HS. Neuroimaging-based biomarkers for treatment selection in major depressive disorder. *Dialogues Clin Neurosci.* (2014) 16:479–90.
78. Drevets WC, Price JL, Simpson JR, Todd RD, Reich T, Vannier M, et al. Subgenual prefrontal cortex abnormalities in mood disorders. *Nature.* (1997) 386:824–7. doi: 10.1038/386824a0
79. Schmaal L, Hibar DP, Sämann PG, Hall GB, Baune BT, Jahanshad N, et al. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the enigma major depressive disorder working group. *Mol Psychiatry.* (2017) 22:900–9. doi: 10.1038/mp.2016.60
80. Mayberg HS. Targeted electrode-based modulation of neural circuits for depression. *J Clin Invest.* (2009) 119:717–25. doi: 10.1172/JCI38454
81. Salustri C, Tecchio F, Zappasodi F, Bevacqua G, Fontana M, Ercolani M, et al. Cortical excitability and rest activity properties in patients with depression. *J Psychiatry Neurosci.* (2007) 32:259–66.
82. Sun P, Wang F, Wang L, Zhang Y, Yamamoto R, Sugai T, et al. Increase in cortical pyramidal cell excitability accompanies depression-like behavior in mice: a transcranial magnetic stimulation study. *J Neurosci.* (2011) 31:16464–72. doi: 10.1523/JNEUROSCI.1542-11.2011
83. De Beurs D, Fried EI, Wetherall K, Cleare S, O' Connor DB, Ferguson E, et al. Exploring the psychology of suicidal ideation: a theory driven network analysis. *Behav Res Ther.* (2019) 120:103419–103419. doi: 10.1016/j.brat.2019.103419
84. O'Connor RC, Kirtley OJ. The integrated motivational-volitional model of suicidal behaviour. *Philos Trans R Soc Lond B Biol Sci.* (2018) 373:20170268. doi: 10.1098/rstb.2017.0268
85. Dwivedi Y, Rizavi HS, Conley RR, Roberts RC, Tamminga CA, Pandey GN. Altered gene expression of brain-derived neurotrophic factor and receptor tyrosine kinase B in postmortem brain of suicide subjects. *Arch Gen Psychiatry.* (2003) 60:804–15. doi: 10.1001/archpsyc.60.8.804
86. Dwivedi Y. Brain-derived neurotrophic factor and suicide pathogenesis. *Ann Med.* (2010) 42:87–96.
87. Berton O, Nestler EJ. New approaches to antidepressant drug discovery: beyond monoamines. *Nat Rev Neurosci.* (2006) 7:137–51. doi: 10.1038/nrn1846
88. Tsankova NM, Berton O, Renthal W, Kumar A, Neve RL, Nestler EJ. Sustained hippocampal chromatin regulation in a mouse model of depression and antidepressant action. *Nat Neurosci.* (2006) 9:519–25. doi: 10.1038/nn1659
89. Wolkowitz OM, Wolf J, Shelly W, Rosser R, Burke HM, Lerner GK, et al. Serum BDNF levels before treatment predict SSRI response in depression. *Prog Neuropsychopharmacol Biol Psychiatry.* (2011) 35:1623–30. doi: 10.1016/j.pnpbp.2011.06.013
90. Holmes S, Scheinost D, Finnema S, Naganawa M, Davis M, DellaGioia N, et al. Lower synaptic density is associated with depression severity and network alterations. *Nat Commun.* (2019) 10:1529–1529. doi: 10.1038/s41467-019-09562-7
91. Mamounas LA, Blue ME, Siuciak JA, Altar CA. Brain-derived neurotrophic factor promotes the survival and sprouting of serotonergic axons in rat brain. *J Neurosci.* (1995) 15:7929–39. doi: 10.1523/JNEUROSCI.15-12-07929.1995
92. Santarelli L, Saxe M, Gross C, Surget A, Battaglia F, Dulawa S, et al. Requirement of hippocampal neurogenesis for the behavioral effects of antidepressants. *Science.* (2003) 301:805–9. doi: 10.1126/science.1083328
93. Berton O, McClung CA, Dileone RJ, Krishnan V, Renthal W, Russo SJ, et al. Essential role of BDNF in the mesolimbic dopamine pathway in social defeat stress. *Science.* (2006) 311:864–8. doi: 10.1126/science.1120972
94. Ramirez-Mahaluf J, Roxin A, Mayberg H, Compte A. A computational model of major depression: the role of glutamate dysfunction on cingulo-frontal network dynamics. *Cereb Cortex.* (2017) 27:660–79.
95. Fee C, Banasr M, Sibille E. Somatostatin-positive gamma-aminobutyric acid interneuron deficits in depression: cortical microcircuit and therapeutic perspectives. *Biol Psychiatry.* (2017) 82:549–59. doi: 10.1016/j.biopsych.2017.05.024
96. Seney ML, Tripp A, McCune S, Lewis DA, Sibille E. Laminar and cellular analyses of reduced somatostatin gene expression in the subgenual anterior cingulate cortex in major depression. *Neurobiol Dis.* (2015) 73:213–9. doi: 10.1016/j.nbd.2014.10.005
97. Yao H, Guet-McCreight A, Mazza F, Moradi Chameh H, Prevot T, Griffiths J, et al. Reduced inhibition in depression impairs stimulus processing in human cortical microcircuits. *bioRxiv.* (2021). [Preprint]. doi: 10.1101/2021.02.17.431698
98. Badcock PB, Davey CG, Whittle S, Allen NB, Friston KJ. The depressed brain: an evolutionary systems theory. *Trends Cogn Sci.* (2017) 21:182–94. doi: 10.1016/j.tics.2017.01.005
99. Corlett PR, Horga G, Fletcher PC, Alderson-Day B, Schmack K, Powers AR. Hallucinations and strong priors. *Trends Cogn Sci.* (2019) 23:114–27. doi: 10.1016/j.tics.2018.12.001

100. Marr D, Poggio T. *From Understanding Computation to Understanding Neural Circuitry*. Cambridge, MA: Massachusetts Institute of Technology, Artificial Intelligence Laboratory (1976).
101. Barack DL, Krakauer JW. Two views on the cognitive brain. *Nat Rev Neurosci*. (2021) 22:359–71. doi: 10.1038/s41583-021-00448-6
102. Brown JC, Higgins ES, George MS. Synaptic plasticity 101: the story of the AMPA receptor for the brain stimulation practitioner. *Neuromodulation*. (2021) 25:1289–98. doi: 10.1016/j.neurom.2021.09.003
103. Yaffe RB, Kerr MS, Damera S, Sarma SV, Inati SK, Zaghoul KA. Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proc Natl Acad Sci USA*. (2014) 111:18727–32. doi: 10.1073/pnas.1417017112
104. Jang AI, Wittig JH, Inati SK, Zaghoul KA. Human cortical neurons in the anterior temporal lobe reinstate spiking activity during verbal memory retrieval. *Curr Biol*. (2017) 27:1700–5.e5. doi: 10.1016/j.cub.2017.05.014
105. Liu TY, Watson BO. Patterned activation of action potential patterns during offline states in the neocortex: replay and non-replay. *Philos Trans R Soc Lond B Biol Sci*. (2020) 375:20190233. doi: 10.1098/rstb.2019.0233
106. Scangos KW, Makhoul GS, Sugrue LP, Chang EF, Krystal AD. State-dependent responses to intracranial brain stimulation in a patient with depression. *Nat Med*. (2021) 27:229–31. doi: 10.1038/s41591-020-01175-8
107. Barbas H, Rempel-Clower N. Cortical structure predicts the pattern of corticocortical connections. *Cereb Cortex*. (1997) 7:635–46. doi: 10.1093/cercor/7.7.635
108. Liu R, Ota K, Duteil S, Duman R, Aghajanian G. Ketamine strengthens CRF-activated amygdala inputs to basal dendrites in mPFC layer V pyramidal cells in the prelimbic but not infralimbic subregion, a key suppressor of stress responses. *Neuropsychopharmacology*. (2015) 40:2066–75. doi: 10.1038/npp.2015.70
109. Opris I, Chang S, Noga BR. What is the evidence for inter-laminar integration in a prefrontal cortical minicolumn? *Front Neuroanat*. (2017) 11:116. doi: 10.3389/fnana.2017.00116
110. Larkum ME, Petro LS, Sachdev RNS, Muckli L. A perspective on cortical layering and layer-spanning neuronal elements. *Front Neuroanat*. (2018) 12:56. doi: 10.3389/fnana.2018.00056
111. Grossberg S, Versace M. Spikes, synchrony, and attentive learning by laminar thalamocortical circuits. *Brain Res*. (2008) 1218:278–312. doi: 10.1016/j.brainres.2008.04.024
112. Wang X, Krystal J. Computational psychiatry. *Neuron*. (2014) 84:638–54. doi: 10.1016/j.neuron.2014.10.018
113. Browning M, Carter C, Chatham C, Den Ouden H, Gillan C, Baker J, et al. Realizing the clinical potential of computational psychiatry: report from the banbury center meeting, February 2019. *Biol Psychiatry*. (2020) 88:e5–10. doi: 10.1016/j.biopsych.2019.12.026
114. Huys Q, Maia T, Frank M. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci*. (2016) 19:404–13. doi: 10.1038/nn.4238