Check for updates

Simulated synapse loss induces depression-like behaviors in deep reinforcement learning

Eric Chalmers[1]*, Santina Duarte[1†], Xena Al-Hejji[1†], Daniel Devoe[1], Aaron Gruber[2‡] and Robert J. McDonald[2‡]

[1]Department of Mathematics and Computing, Mount Royal University, Calgary, AB, Canada, [2]Canadian Centre for Behavioural Neuroscience, University of Lethbridge, Lethbridge, AB, Canada

Deep Reinforcement Learning is a branch of artificial intelligence that uses artificial neural networks to model reward-based learning as it occurs in biological agents. Here we modify a Deep Reinforcement Learning approach by imposing a suppressive effect on the connections between neurons in the artificial network—simulating the effect of dendritic spine loss as observed in major depressive disorder (MDD). Surprisingly, this simulated spine loss is sufficient to induce a variety of MDD-like behaviors in the artificially intelligent agent, including anhedonia, increased temporal discounting, avoidance, and an altered exploration/exploitation balance. Furthermore, simulating alternative and longstanding reward-processing-centric conceptions of MDD (dysfunction of the dopamine system, altered reward discounting, context-dependent learning rates, increased exploration) does not produce the same range of MDD-like behaviors. These results support a conceptual model of MDD as a reduction of brain connectivity (and thus information-processing capacity) rather than an imbalance in monoamines—though the computational model suggests a possible explanation for the dysfunction of dopamine systems in MDD. Reversing the spine-loss effect in our computational MDD model can lead to rescue of rewarding behavior under some conditions. This supports the search for treatments that increase plasticity and synaptogenesis, and the model suggests some implications for their effective administration.

Introduction

Reinforcement learning: reward-based learning and behavior

Reinforcement Learning (RL) is a field of science that provides mathematical models of reward-based learning and decision-making. With its origins in the disparate psychological and computational fields of trial-and-error learning and optimal control (Sutton and Barto, 2018), it is now a field where artificial intelligence (AI) and brain sciences overlap: recent RL-based learning algorithms can play *Go* and *Starcraft* at superhuman levels (Schrittwieser et al., 2020; Vinyals et al., 2019), while psychologists and neuroscientists use RL models to understand animal behavior (Daw, 2012; Neftci and Averbeck, 2019; Sutton and Barto, 2018).

Temporal difference learning is an important branch of RL which models trial-and-error learning. Temporal difference learning measures the difference between expected and actual rewards (the "temporal difference" or "reward prediction error"; RPE) each time the agent executes an action. The estimated value of that action is then adjusted in proportion to the

RPE, according to a learning rate parameter. In the 1990s it was hypothesized that this model actually corresponds to learning mechanisms in the dopaminergic and striatal systems (Montague et al., 1996; Schultz et al., 1997). This hypothesis has grown steadily, and today it is generally thought that the phasic activity of dopamine neurons in the midbrain signals reward prediction error (RPE) (Montague et al., 1996). It is believed that these phasic activity fluctuations influence plasticity in the striatum (Schultz et al., 1997) (and possibly hippocampus (Mehrotra and Dubé, 2023) and prefrontal cortex (Daw et al., 2005)) in a way that optimizes the organism's expected or perceived value of particular actions, allowing it to learn rewarding behavior.

Deep Reinforcement Learning is a recent branch of RL which incorporates artificial neural networks: the networks are embedded within and control the actions of an artificial agent. In one possible implementation, the network accepts sensory input describing the environmental state, and the resulting activations of its output neurons represent the estimated values of various actions, given that state. Reward prediction errors experienced as the agent interacts with its environment are used to adjust the network's internal connections in a learning process that improves the quality of the value estimates (thus maximizing the agent's ability to make rewarding decisions), and generalizes knowledge between similar states (a critical ingredient in the ability to solve large problems and learn complex behaviors).

Because of the close analogy between Deep RL and biological reward-based learning, Deep RL is starting to be used in significant neuroscientific modeling and hypothesis creation, though much potential remains untapped (Botvinick et al., 2020). In particular, there is a tremendous but little-explored opportunity to use Deep RL to model *dysfunctions* of learning and behavior, including those associated with Major Depressive Disorder (MDD) (Mukherjee et al., 2023). Artificial intelligence technologies have been used in diagnosis and treatment of depression (Lee et al., 2021), but a computational model of the disorder itself would be immensely valuable for MDD research: such models are cheaper and more ethical than animal models, can provide new ways of conceptualizing the disorder, and can produce new hypotheses for experimental researchers to investigate. This paper presents just such a model.

Major depressive disorder—through a reinforcement learning lens

Major Depressive Disorder (MDD) is a complex psychological disorder. It is both highly disabling and highly prevalent: according to the National Institute of Mental Health, over 21 million Americans over the age of 18 have experienced a Major Depressive Episode over the last year (National Institute of Mental Health, n.d.). According to the DSM-5, symptoms of MDD include anhedonia, withdrawal from social activities, loss of interest, and increased indecisiveness (American Psychiatric Association, 2022). The ICD-10 categorizes recurrent depressive disorder as repeated episodes of depressive symptoms, including a decrease in overall mood and a reduction in the capacity for enjoyment of personal interests (World Health Organization, 1993). Overall, MDD involves a general decrease in goal-directed behavior and productivity.

MDD is complex, heterogeneous, and not completely understood, but has been conceptualized in various ways. As it happens, some prominent conceptions involve or overlap with Reinforcement Learning and could be modeled using RL or straightforward extensions of RL. These conceptions of MDD include:

Monoamine deficiency

The "monoamine hypothesis"—the idea that MDD is primarily a dysregulation of monoaminergic systems (dopamine, serotonin, etc.)—has been used for many years and inspired a variety of antidepressant drugs. Early generations of tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs) eventually gave way to selective serotonin reuptake inhibitors (SSRIs), which can take several weeks to work but have fewer side effects (Edinoff et al., 2021). Amphetamines have sometimes been used (Tremblay et al., 2002) or paired with other drugs (Pary et al., 2015) because they increase dopamine levels. The downregulation of dopaminergic systems in MDD has been well documented (Belujon and Grace, 2017) and is of particular interest from a Reinforcement Learning perspective: if dopamine signals reward prediction error, then disruption of that signal may also disrupt learning.

Increased temporal discounting

Temporal or delay discounting refers to the decrease in the perceived value of a future reward as the time it takes to reach the reward increases. Higher discounting rates mean a stronger preference for immediate reward over a delayed reward. It has been hypothesized that tonic activity of dopamine neurons affects discounting (Pattij and Vanderschuren, 2008; Smith et al., 2005), and an increase in delay discounting has been associated with various psychiatric disorders, including bipolar disorder, bulimia nervosa, borderline personality disorder, and ADHD (Amlung et al., 2019; Maia and Frank, 2011), and larger discounting rates are observed in depressed individuals relative to controls (Imhoff et al., 2014; Pulcu et al., 2014). This altered discounting may be a reflection of particular symptoms of MDD, such as feelings of hopelessness (Pulcu et al., 2014), or related to the impaired episodic future thinking that is observed in depressed individuals (Hallford et al., 2020). While MDD is generally associated with increased discounting, it should be noted that the discounting phenomenon is complex and may vary with age and environmental factors (Lempert and Pizzagalli, 2010; Li et al., 2012; Read and Read, 2004; Whelan and McHugh, 2009).

Faster learning from negative experiences

Depressed individuals appear to have a heightened response to punishment or negative experiences and a blunted response to reward or positive experiences, both neurally and behaviorally (Eshel and Roiser, 2010; Maddox et al., 2012; Rygula et al., 2018). Such imbalances in feedback processing have been linked to altered dopamine signaling (Sojitra et al., 2018) and are also affected by SSRIs (Herzallah et al., 2013). Anhedonic severity seems to correlate with attenuated processing of positive information in the striatum and enhanced performance in avoiding losses (Reinen et al., 2021). However, reinforcement learning models with dual learning rates (separate learning rates for rewards and punishments) have not always detected faster learning from punishments when fit to behavioral data (Brolsma et al., 2022; Pike and Robinson, 2022; Vandendriessche et al., 2023).

Altered exploration/exploitation balance

MDD is associated with altered exploration-exploitation tradeoff: the problem of deciding whether to exploit a trusted, rewarding

option or explore other options that may potentially yield better outcomes. Depressed individuals exhibit an increase in stochastic choices and exploratory behavior (Blanco et al., 2013; Kunisato et al., 2012; Rupprechter et al., 2018), and this has been interpreted as the result of a decreased sensitivity to the value of different options (Chung et al., 2017; Huys et al., 2013; Mukherjee et al., 2023). However, as with the impaired processing of negative versus positive feedback, some studies that apply computational RL models have found that exploration levels do not account for the differences between MDD patients and controls (Bakic et al., 2017; Rothkirch et al., 2017).

A new model: depression as reduced brain connectivity

Depression has been shown to affect dendritic spine density. Depression is generally associated with decreased spine density in the hippocampus and prefrontal cortex and increased density in the amygdala and nucleus accumbens (Helm et al., 2018; Qiao et al., 2016; Runge et al., 2020). Changes in spine density may depend on brain region, sex, and other experimental conditions (Qiao et al., 2016). Overall, lower synaptic density correlates with depression severity (Holmes et al., 2019) and the apparent loss of spines in the prefrontal cortex and hippocampus is a particular focus of this paper.
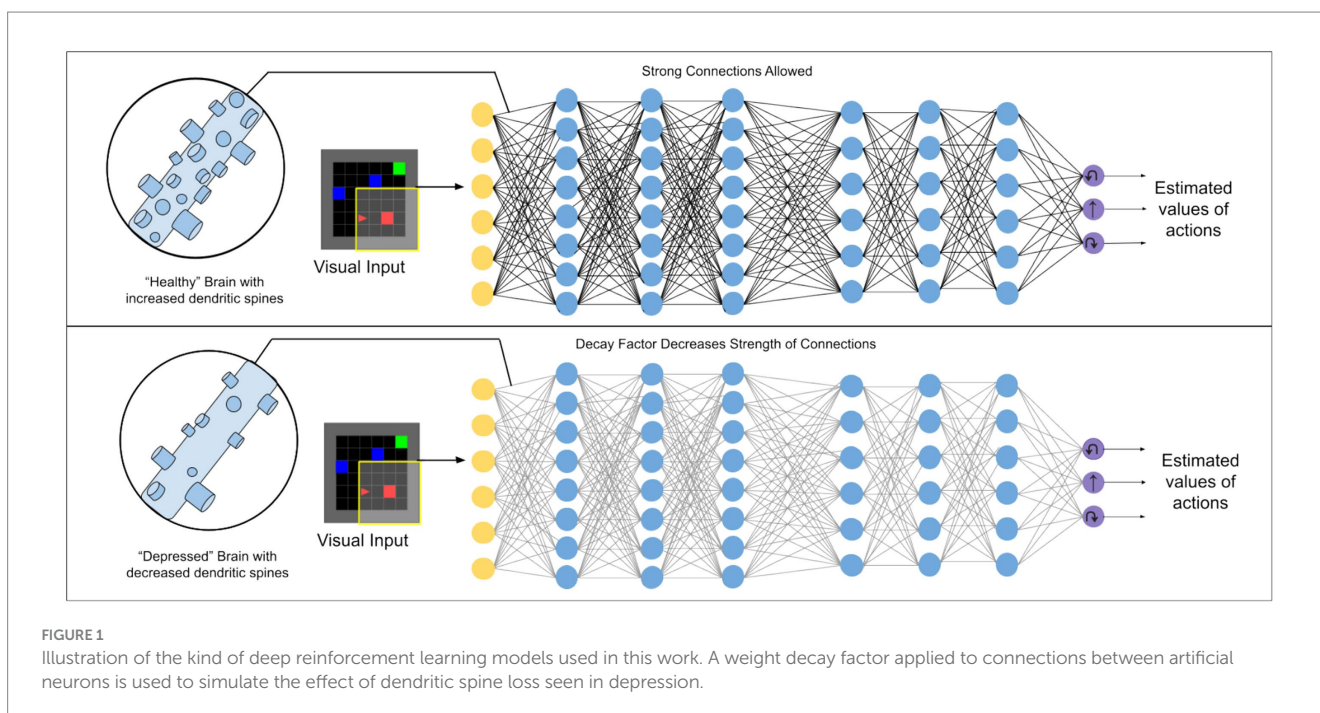
Dendritic spines are small protrusions from neuronal dendrites that can contact and form a synapse with a neighboring neuron. Fewer spines therefore suggests reduced overall brain connectivity. Some recent work suggests that ketamine relieves depression by promoting plasticity and spinogenesis (Krystal et al., 2019, 2023) and that spines formed after ketamine treatment are necessary for sustained antidepressant effects (Moda-Sava et al., 2019). Furthermore, it has been suggested that conventional antidepressants (that are based largely on the monoamine hypothesis) may work not by modulating monoamine concentrations *per se*, but because they promote slow synaptogenesis (Johansen et al., 2023; Speranza et al., 2017; Wong et al., 2017), or because they promote cognitive reconsolidation in a more positive way (Harmer et al., 2009a, 2009b). Observations like these have affected a recent shift in thinking, away from MDD as a monoamine system dysfunction, toward seeing MDD as a dysfunction of neuroplasticity (Liu et al., 2017).

Towards a computational model of depression

In this paper we set up a deep reinforcement learning algorithm to simulate dendritic spine loss. Specifically, we apply a "weight decay" effect to all connections between artificial neurons, causing each to degenerate over time at a rate proportional to their strength (i.e., the strongest communication channels are affected most—see Methods). This simulates the effect of spine loss by impairing the connections between neurons (illustrated in Figure 1).

We show, perhaps surprisingly, that this simulated spine loss is sufficient to induce a range of depression-like behaviors in the artificial agent, including anhedonia, temporal discounting, avoidance, and increased exploration. We further demonstrate that simulating alternate RL-centric conceptions of depression (reduced dopamine signaling, an artificially-increased discount rate, faster learning from negative experience, and altered exploration/ exploitation balance) do not produce the same range of behaviors, showing that the simulated spine-loss model has face validity as a model of depression. We conclude by drawing several interesting hypotheses and implications from the model, including a new way to conceptualize depression, a speculative explanation for dopamine system dysregulation in MDD, and implications for treatments.

FIGURE 1
Illustration of the kind of deep reinforcement learning models used in this work. A weight decay factor applied to connections between artificial neurons is used to simulate the effect of dendritic spine loss seen in depression.

Results

A simulated "world"

Experiments in this paper use a simulated goal-seeking task illustrated in Figure 2a. The simulation places an agent in a small room which it must learn to navigate. Three types of objects are present in the room: several "optional" goals, one "required" goal, and a "hazard." Optional goals appear randomly throughout the room in each episode, and deliver a small reward to the agent when collected. The required goal always appears in the same location. When it is reached, the agent receives a large reward, the current episode ends, and the next begins. The hazard delivers a punishment (negative reward) to the agent when touched. At each step the agent chooses one of three actions: turn left, turn right, or move forward. The agent's visual field only covers part of the room, so it must learn how to select appropriate actions to maximize reward, given the incomplete visual information. Both "healthy" and "depressed" agents (with simulated spine loss) were placed in this environment and allowed to learn behavioral strategies over a sufficient period of time.

While this simulation presents a goal-seeking problem involving spatial navigation, it should also be seen as a simple metaphor for the sequential decision-making of daily life. The optional goals represent opportunities like play, exploring interests, or socialization, while the required goal represents basic survival strategies like finding food or employment, which must be attended to every day.

Simulated spine loss induces depression-like behaviors

Anhedonia

Anhedonia (loss of interest or pleasure) is recognized as a hallmark symptom of MDD by both the DSM-5 and the ICD-10 (American Psychiatric Association, 2022; World Health Organization, 1993). Behavior of both healthy and spine-loss agents is illustrated in Figure 2. Comparing the times required to reach asymptotic slope in the cumulative reward curves and comparing the asymptotic slopes themselves, we see that the "depressed" agents learn a simple strategy more quickly than the "healthy" agents, but the strategy learned by the
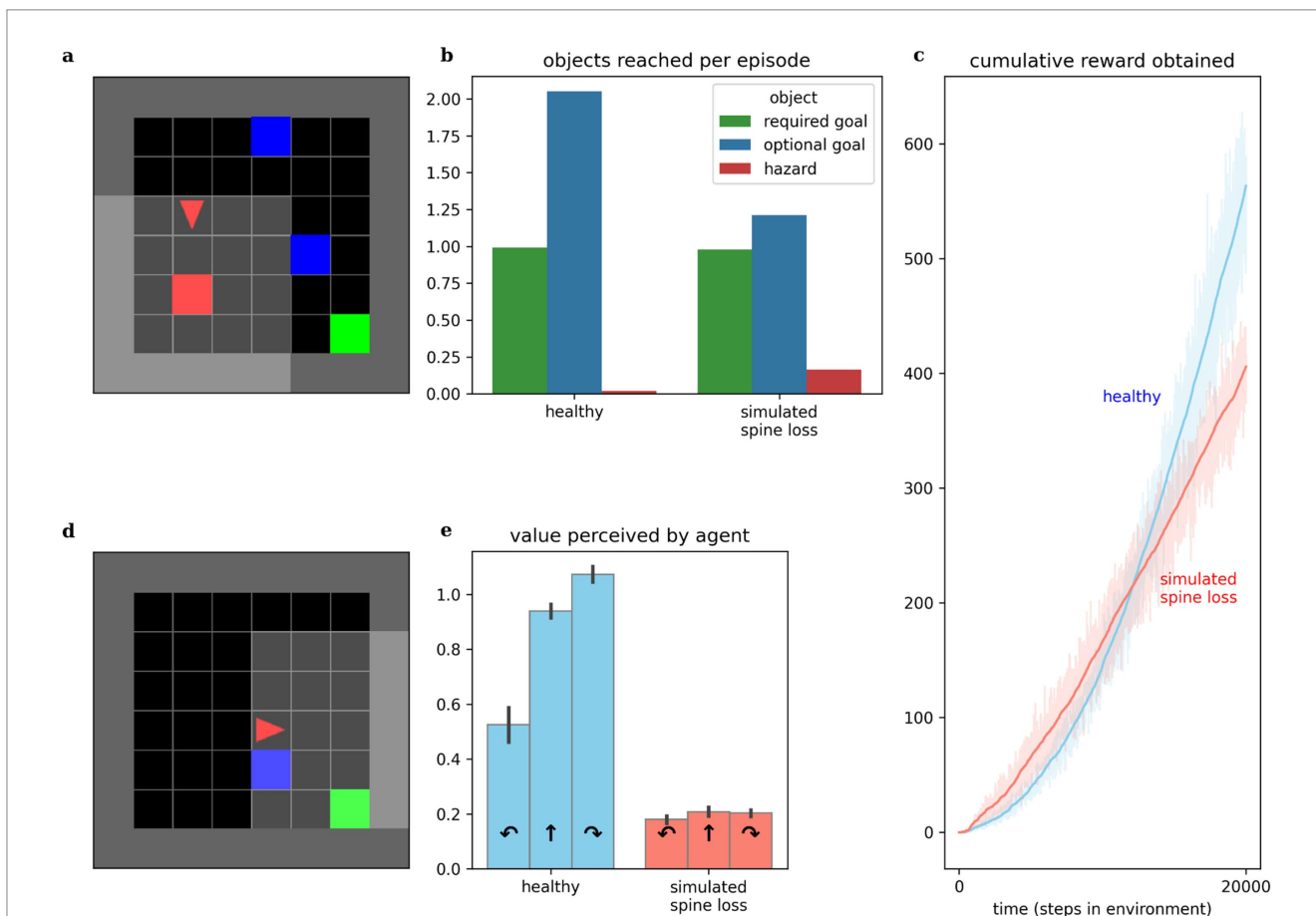
FIGURE 2
Comparing behaviors of simulated healthy and simulated spine loss agents. (a) The simulated "world." The agent (red triangle) must learn to navigate the room in search of the green goal. Blue boxes are optional bonus rewards that can be collected en route to the goal, and red boxes are hazards that bring a negative reward (punishment). (b) The simulated spine loss agent still reaches the goal in each episode but collects fewer bonus rewards. (c) The simulated spine loss agent learns a less-rewarding strategy than the healthy agent but learns it quicker, as seen by the time to arrive at asymptotic slope. (d) Contrived situation in which the agent may bypass the optional reward or take an extra step to collect it en route to the goal (reward optimal). (e) Agents' perceived values for the contrived situation in (d)—the healthy agent's perceived value of detouring through the optional reward is very high. The spine loss agent has much weaker preferences and a slight preference for bypassing the optional reward (an anhedonia-like effect). Errorbars and shaded regions show the 95% confidence interval of the mean over 20 repetitions.

"healthy" agents is more rewarding (see Figure 2c). This is illustrated further by comparing the number of optional goals obtained per episode. "Healthy" agents collect almost two optional goals per episode, while "depressed" agents tend to collect only 1. If optional goals represent opportunities like play, exploring interests, or socialization, then these results mirror the loss of interest in these opportunities that accompanies depression.

Studies of MDD in animals often detect anhedonic (reduced pleasure-seeking) behavior using a sucrose preference test (Bessa et al., 2013; Zhuang et al., 2019). Here, we instead place agents in the contrived scenario illustrated in Figure 2d and observe their internal, perceived values. Assuming reasonably low reward discounting, the reward-optimal strategy would turn right to collect the optional goal en route to the required goal—requiring 1 additional action but

obtaining both goals. The depressed agent chooses to bypass the optional goal—assigning a higher value to moving forward than to turning right. This is consistent with the results in Figure 2b and suggests a simulated anhedonia. But the depressed agent also shows perceived values that are reduced categorically, revealing a general loss of interest underlying the anhedonic behavior.

Increased discounting

Altered discounting has been observed in MDD, both monetarily (Pulcu et al., 2014) and in response to basic rewards through simulations (Rupprechter et al., 2018). Computational models provide the opportunity to directly expose the agent's internal preferences and perceived values. We can infer an agent's effective temporal discounting rate from these values. Figure 3b shows the simulated
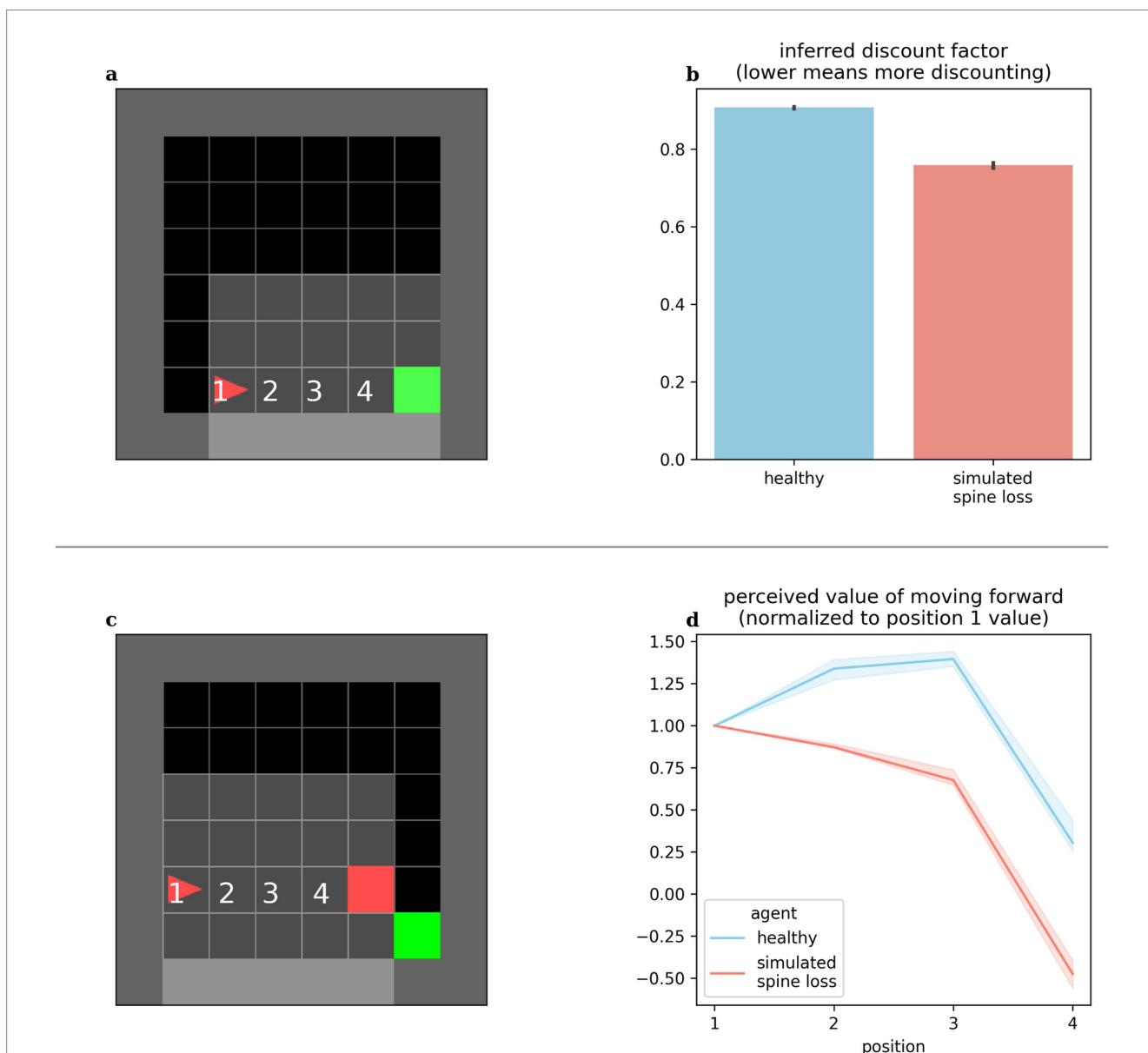
FIGURE 3
(a,b) Agents' effective discounting rates can be inferred by placing them progressively closer to the goal and measuring their perceived values. The spine loss agent operates with a lower effective discount factor (more discounting). (c,d) Contrived situation in which the agent is moved toward a hazard. The healthy agent's perceived value of moving forward increases through positions 1−3 (moving forward from these positions brings the goal closer). Only in position 4 does the healthy agent's perceived value of moving forward drop. For the spine loss agent this drop is generalized inappropriately to positions 2 and 3. Errorbars and shaded regions show the 95% confidence interval of the mean over 20 repetitions.

spine loss agent operating with a lower effective discount factor (more reward discounting). Surprisingly, this occurs *even though* all RL algorithms in these experiments used the same explicit discount factor setting. That is, the spine loss seems to induce an additional discounting effect.

Avoidance

MDD involves maladaptive and increased avoidance behavior (Ironside et al., 2020; Ottenbreit et al., 2014). We placed our computational agents in the contrived situation shown in Figure 3c: the agent is slowly moved toward a hazard lying in front of the required goal. The healthy agent's perceived value for forward motion increased through positions 1–3 (since this brings the agent closer to the goal). Only in position 4 does the value of moving forward drop—at that point the agent would prefer to go around the hazard. Conversely, in the depressed agent, the aversion to moving forward is (inappropriately) generalized to positions 2 and 3. Broekens et al. have suggested that such a decrease in perceived value within an RL model can be interpreted as analogous to fear (Broekens et al., 2015).

Simulated spine density modulates depression

By applying and then removing the dropout effect in the agent's neural network, we can simulate the loss of dendritic spines and their subsequent restoration. Results in Figure 4 show that depressive behaviors can be alternately induced and removed by applying or removing the connection weight decay factor. This seems to suggest that spine density, rather than being simply *correlated* with depression, modulates it. Interestingly, removing the simulated spine loss allows a return to the agent's original behavior, but only after a brief drop in performance and a period of re-learning (this period is discussed further in the Discussion section).

Simulated spine loss affects network learning and information content

Observing changes in the artificial neural networks over time, we see that a given reward prediction error generally induces less change in the network with simulated spine loss. Figure 5 shows a dramatic change in spine-loss-network weights early in learning (Figure 2c suggests this represents the network quickly learning a basic survival strategy) but thereafter the spine-loss network shows less plasticity than the "healthy" network. Additionally, when the agents move toward the required goal, most neurons in the spine-loss network show activity which is highly correlated with this movement, suggesting that most of the network is dedicated to information about the required goal. In the healthy network, some neurons strongly correlate with the goal-directed movement, but others are not—suggesting those neurons are more concerned with other things (presumably information related to the optional goals and the hazard).

Simpler learning tasks are not impaired

The results above may give the impression that the spine-loss agent is *generally* impaired relative to the "healthy" agent. But Figures 2b,c show the "depressed" agent converging on a strategy faster than the "healthy" agent—a strategy that does allow it to reach the goal every episode. To investigate this further, we create a simplified version of the environment by removing the optional rewards and hazards, leaving only the required goal. In this simpler task the spine-loss agent actually performs slightly better than the control agent (see Figure 6). This seems to suggest that spine loss has a greater effect on more complex behaviors that require higher-order processing. One interpretation is that spine loss reduces overall network capacity, but leaves enough capacity for the network to learn simpler behaviors (see the Discussion section for further commentary).

FIGURE 4
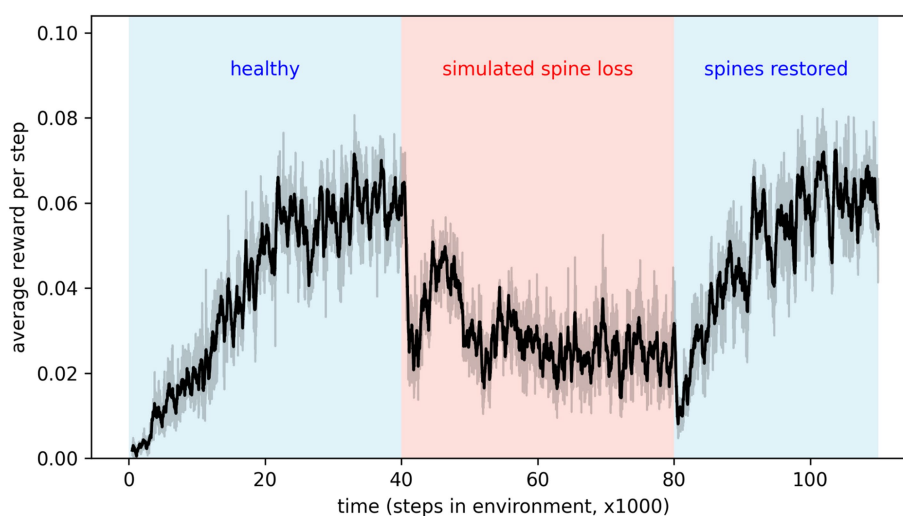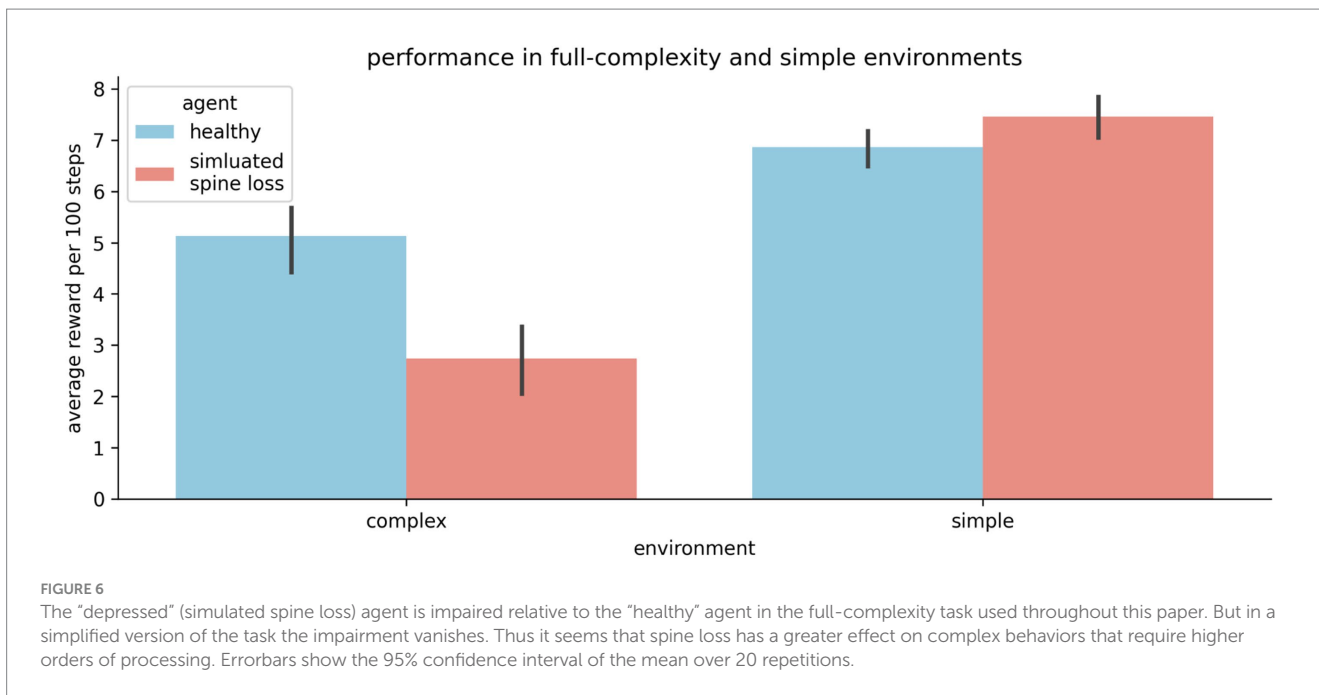Applying simulated spine loss to a healthy agent causes it to revert to the simpler, low-reward depression-like behavior. Relieving the spine loss (restoring the spines) allows a return to the original behavior—after a short readjustment period. This may support the idea that spine density modulates depressed cognition and behavior. Shaded regions show the 95% confidence interval of the mean over 20 repetitions.

FIGURE 5
(a) Weight changes induced in the networks per unit loss. This measures simulated response to the reward prediction error signal (i.e., delivered by dopamine). The spine loss network experiences dramatic alterations early in learning, allowing it to converge on a basic strategy quickly. The healthy network exhibits greater and sustained plasticity throughout learning. This effect may hint at an explanation for dopamine system dysfunction in depression. (b) Most neurons in the network with simulated spine loss have activations highly correlated with proximity to the goal, indicating that almost the entire network has been used to store information related to the basic goal-seeking strategy. Neurons in the healthy agent's network may store a greater variety of information. Shaded regions and boxplots show the 95% confidence interval of the mean over 20 repetitions.

FIGURE 6
The "depressed" (simulated spine loss) agent is impaired relative to the "healthy" agent in the full-complexity task used throughout this paper. But in a simplified version of the task the impairment vanishes. Thus it seems that spine loss has a greater effect on complex behaviors that require higher orders of processing. Errorbars show the 95% confidence interval of the mean over 20 repetitions.

A small amount of simulated spine loss is beneficial

The relationship between spine density and cognition / behavior is not well understood. And while depression is associated with decreased spine density, spines and synapses can be lost for many other reasons. During developmental pruning, for example, many synapses are eliminated from a young developing brain. Here we model adult brains in which the synaptic pruning process has

already completed, but even in adulthood there is normal, ongoing turnover of a small number of spines and synapses (Runge et al., 2020).

Some turnover of connections is probably good, preventing stagnation and overfitting. This is well understood by machine learning practitioners, who often apply a small amount of weight decay to their artificial neural networks to improve learning and generalization (Andriushchenko et al., 2023). To illustrate this, we sweep through a range of weight decay settings from mild (representing normal, ongoing turnover of synapses) to extreme (representing pathological spine loss).

As shown in Figure 7, weight decay improves the agent's performance up to a point: large amounts of decay cause depression-like impairments.

Other conceptual models do not produce the same depression-like behaviors when simulated

We simulate the four other conceptions of MDD listed in the introduction: dopamine deficiency, increased temporal discounting, faster learning from negative experiences, and altered exploration/exploitation balance. Surprisingly, none produce all the same depression-like behaviors as simulated spine loss (Table 1).

Discussion

Implications for our conception of MDD

Computational models can help us fine-tune our conceptualization of a disorder. This computational model simulates spine density loss in a neural network, and surprisingly we find that this is sufficient to produce a variety of depression-like phenomena. These results encourage us to supplement our current understanding of MDD with the new (and complementary) view of MDD as a reversible loss of brain capacity.

If neural networks store information in connection weights or synapses (Kozachkov et al., 2022), then, all other things being equal, a brain with more connections can store more information. Spine density loss reduces the number of connections available to the neural network's learning process and thus reduces the capacity for storing information that supports choice strategies and behaviors. Thus, the

reversion or reduction to the simpler, basic survival strategy when spine loss is applied in Figure 4: there is no longer enough room to store the richer, more rewarding, but more complex strategy. Similarly, if a depressed human brain only has enough room to hold the necessary behavior of going to work, then hobbies, socialization and other interests could be squeezed out of the neural network to some degree. This interpretation adds to (and seems roughly consistent with) the recent shift in thought away from MDD as a monoamine system dysfunction toward the idea of MDD as a dysfunction of neuroplasticity (Liu et al., 2017).

Future experimental studies could explore this idea by correlating neural activity with task performance in behaving animals. If the task is complex, such experiments should expect to see reduced task performance but clearer neural encoding of the main features of the task in depressed animals—a few results along these lines have been observed in other work (Gruber et al., 2010), and would be consistent with our Figure 5.

The idea of MDD as loss of brain capacity may also begin to reconcile some conflicting observations from computational literature. Implementing dual learning rates in reinforcement learning models has had mixed success in explaining depressed behaviors (Brolsma et al., 2022; Pike and Robinson, 2022; Vandendriessche et al., 2023). Similarly, fitting exploration/exploitation parameters in computational models has had mixed success in accounting for depressed behavior (Bakic et al., 2017; Pike and Robinson, 2022). Our results may provide a clue for interpreting these findings. Our agents all used the same explicit temporal discount factor settings, yet the agent with simulated spine loss demonstrated additional discounting. That is, the reduced neural-network connectivity creates a discounting effect completely independent of the discount parameters used in conventional RL models—probably because the network must focus its limited capacity on learning how to seize nearby rewards. Similarly, faster learning from negative experiences and increased exploration

FIGURE 7
The relationship between weight decay and performance in our simulations. A small amount of weight decay (representing the normal, ongoing turnover of spines and synapses) is beneficial. More extreme weight decay causes impairment. Shaded region shows the 95% confidence interval of the mean over 20 repetitions.
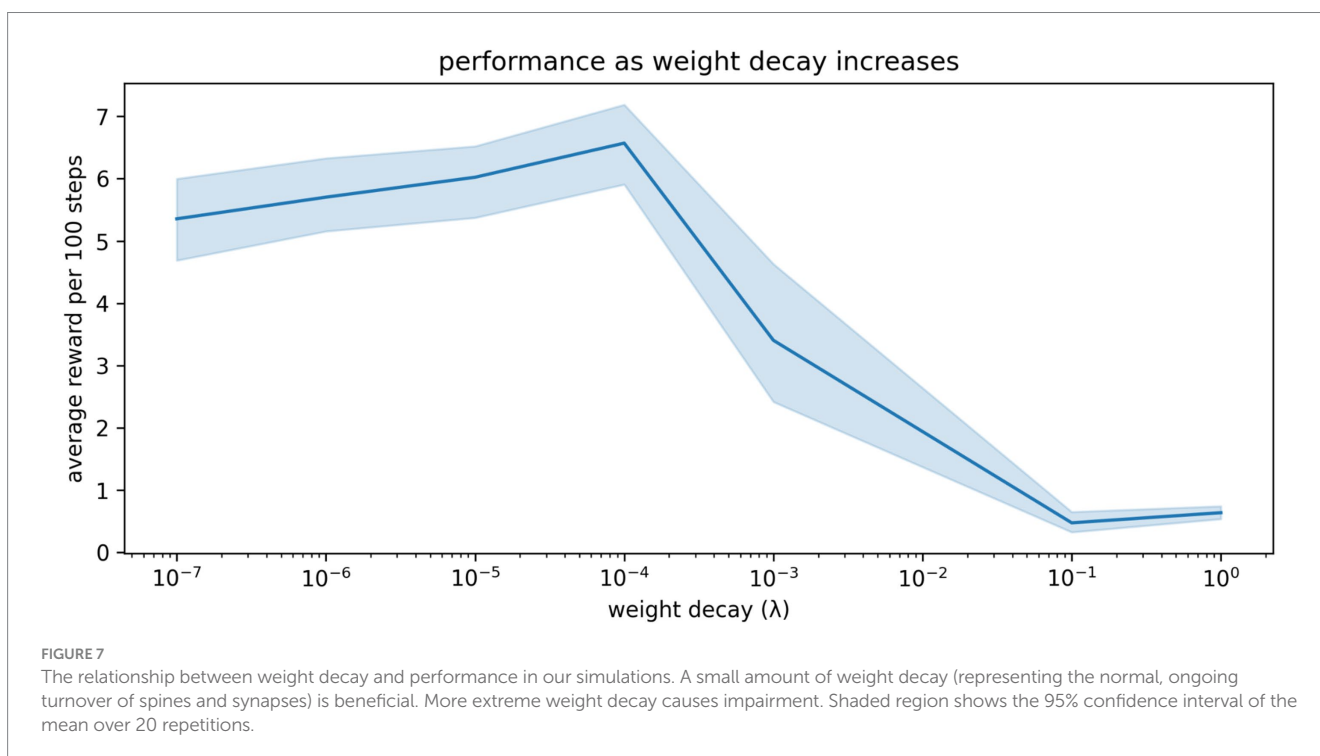
TABLE 1 The spine-loss model produces a wider range of depression-like behaviors than alternative conceptions of depression.

		Expected behavior			
		Anhedonia	Generalized avoidance/fear	Increased discounting	More exploration
Model of MDD	Spine loss	✓	✓	✓	✓
	Reduced reward-prediction-error signal				
	Faster learning from negative experiences				
	Higher discount factor			✓	✓
	Increased exploration				✓

may arise from reduced brain connectivity—that is, they may be secondary effects in the same way that increased temporal discounting was a side-effect of reduced connectivity in our model. In that case they would not be accurately accounted for by the explicit parameters used in conventional RL models. Future experimental studies could verify this conclusion by fitting our new deep-RL model to behavioral data, to test whether the depressed behaviors can be accounted for by our model without explicit dual-learning rates or exploration parameters.

Interestingly, while spine density has been observed to decrease in the prefrontal cortex during MDD, Smoski et al. observed an *increase* in responsivity in the orbitofrontal cortex connected to risk and reward processing (Smoski et al., 2009). Our model – which models spine loss in the prefrontal cortex, also provides an account of this increased responsivity. In addition to the weights that represent communication channels with neighboring neurons, each neuron in an artificial neural network has a bias parameter that represents the sensitivity of its firing threshold, and so is related to its responsiveness in general. These biases are allowed to change during the training process just as the network's weights do (see Methods). When we apply large weight decay to the weights to simulate spine loss, we observe that the network attempts to compensate by increasing neuron biases—making them more sensitive. This increase is likely the neurons' attempt to compensate for the net reduction in input signal caused by the simulated spine loss. See Supplementary materials for more details.

Dopamine system dysregulation as a cause of MDD? or an effect?

Surprisingly, simulating spine loss in an artificial neural network is sufficient to produce a range of depression-like behaviors, but simulating reduced reward-prediction-error signaling by dopamine apparently is not. Still, dopamine system dysregulation is a well-documented feature of depression (Belujon and Grace, 2017). How do we reconcile these facts?

The computational model suggests a speculative explanation for dopamine system dysfunction. Spine loss and reduced brain connectivity would limit the information-storage capacity of a neural network (Kozachkov et al., 2022; Mayford et al., 2012). Even basic survival strategies must fight for representation in a network with limited capacity, so there is less freedom for the network to adapt to the environment in arbitrary ways. Under these conditions, a dopamine spike unrelated to survival-critical events may go "unanswered" in terms of network changes. This explains the

results in Figure 5, which show that the spine-loss-network's response to reward prediction error [i.e., the dopamine signal (Montague et al., 1996; Schultz et al., 1997)] is muted relative to the healthy network. With fewer opportunities to act, the dopamine system may be thrown into dysregulation. Under this view, dopamine system dysregulation is not a *cause* of MDD; rather, spine loss and reduced brain connectivity cause *both* MDD and dopamine system dysregulation. It has been proposed that early damage to the hippocampus may lead to the altered forebrain dopamine operation seen in Schizophrenia (Hanlon and Sutherland, 2000)—perhaps the altered dopamine operation in MDD is, similarly, secondary.

There is a chance that thinking about MDD as a reversible loss of brain capacity could allow a wider (and very ambitious) reconciliation between the monoamine hypothesis and the newer neuroplasticity theory of MDD, by seeing neural network capacity as the link between the two. If the activation of 5-HT7 receptors promotes the formation of dendritic spines (Speranza et al., 2017), and if those spines prevent depression-like symptoms by increasing network capacity (Johansen et al., 2023), this may partially account for serotonin and tryptophan depletion causing depression-like symptoms in patients (Jauhar et al., 2023)—the type of observation that originally led to the widespread use of SSRIs under the monoamine hypothesis (Liu et al., 2017). Clearly this is an inadequate summary of a complex [and contentious (Jauhar et al., 2023; Moncrieff et al., 2023)] topic. Still, limiting network capacity seems to account for a surprising range of depression-like behaviors—suggesting that various systems and interventions may exert their influence on depressive states *through* network capacity.

Implications for MDD treatments

The model makes the hypothesis that spine density directly modulates depression. Many studies have associated spine density loss with depression, and some evidence for a causal link has been provided in the context of ketamine treatment (Moda-Sava et al., 2019). But in a recent review of psychedelic effects on neuroplasticity, Calder & Hasler conclude, "though changes in neuroplasticity and changes in cognition or behavior may occur simultaneously, whether neuroplasticity mediated those changes remains an open question for future studies to address" (Calder and Hasler, 2023). In light of this, the results in Figure 4 are especially significant since they show the simulated spine-loss effect directly modulating depression-like behaviors. Though speculative, these results strengthen the argument

for increased neuroplasticity and spine density as the sources of improved cognition and behavior, and support the search for treatments that promote spine growth.

This hypothesis that spine density modulates depression is consistent with the new thinking about MDD as a dysfunction of neuroplasticity, which has prompted research into treatments that upregulate plasticity. For example, the N-methyl-d-aspartate (NMDA) receptor antagonist, Ketamine, is theorized to promote synaptogenesis through one of several possible pathways that increase brain-derived neurotrophic factor (BDNF) levels and, ultimately, synaptic plasticity (Castrén and Monteggia, 2021; Duman et al., 2016; Krystal et al., 2019; Qiao et al., 2016; Zhou et al., 2014). Ketamine has indeed been observed to increase BDNF levels (Yang et al., 2013), increase spinogenesis and reverse dendritic atrophy (Li et al., 2010, 2011, p. 20; Moda-Sava et al., 2019), and produce fast and persistent antidepressant effects in patients and animal models (Berman et al., 2000; Browne and Lucki, 2013; Krystal et al., 2019; Murrough et al., 2013; Zarate et al., 2006). Psychedelics such as D-lysergic acid diethylamide (LSD), psilocybin, and dimethyltryptamine (DMT) have also been shown to promote neuroplasticity—possibly through a pathway that starts with serotonin 5-HT2A receptor agonism and involves stimulating BDNF production (Aleksandrova and Phillips, 2021; Calder and Hasler, 2023; Ly et al., 2018). They too can induce rapid antidepressant effects (dos Santos et al., 2016; Romeo et al., 2020; Rucker et al., 2016).

Furthermore, our model suggests some implications for the administration of such treatments. In Figure 4, the period immediately after the simulated spine restoration is interesting because the agent does not immediately reacquire the high-reward behavior. Instead, the rate of reward actually drops, then slowly climbs back to pre-depression levels as the agent undergoes *general re-learning* of rewarding behavior for the environment. During this period the network is not simply augmenting the basic survival strategy stored during depression with new information about optional goals. Rather, the entire network is being reconfigured to make use of the new connections.

If a biological agent's brain undergoes a similar (if much less dramatic) experience, then this period of time immediately after restoring lost spines is critical: experiences during this phase will influence what is re-learned. It is already known in the context of psychedelic treatments, for example, that the supportiveness of the environment during and after administration affects the treatment outcome (Calder and Hasler, 2023). More research may be necessary to determine how to best use the "window of neuroplasticity" that remains open for a short time after treatment (Calder and Hasler, 2023). We note that our artificial agent would not have reacquired the high-reward behavior if placed in an empty room after spine restoration: it requires the right environmental exposure to re-learn an appropriate strategy. This is similar to the conclusion drawn by Harmer et al. regarding conventional antidepressants: that they take effect slowly because they increase positive emotional processing, after which the patient must gradually re-learn their relationship to the world through accumulation of positively-processed experience (Harmer et al., 2009a,b).

More research is needed into the "window of neuroplasticity" effect generally, and based on these results, future experimental studies might examine the effects of environment and experience shortly after a plasticity-inducing treatment is administered.

Limitations of the computational model, and opportunities for future work

As scientists, we often understand complex biological systems through *models* (e.g., animal models, mathematical / computational models, block diagrams, pictorial representations, etc.). Every model is a sort of metaphor: it is *not* the complex target system, but it has something in common with—and therefore tells us something about—that system. Like all good models, our computational model abstracts away complexity in order to highlight some general principles. A real brain involves modularity, sparse and small-world connectivity, interactions between distinct structures, cortical columns, polysynaptic pathways, etc. Conventional artificial neural networks abstract away much of this complexity, and highlight the basic principle that rewarding behavior is learned through alteration of neuron-to-neuron connections. In reality, each of those neuron-to-neuron connections consists of a set of (sometimes redundant) synapses (Hiratani and Fukai, 2018) that develop from among a large number of filopodia (Runge et al., 2020) across various functional zones of a dendrite (Hawkins and Ahmad, 2016). An artificial network abstracts away much of this complexity too, using a single weight to represent the *net* communication channel between two neurons. Such abstractions are allowing researchers to build tractable models of neural reinforcement learning; useful for testing high-level ideas and generating hypotheses (Botvinick et al., 2020).

Spine loss suggests fewer synapses and overall reduced connectivity between neurons. In an artificial neural network (where the usual approach is to represent the net connection between two neurons using a single weight) applying weight decay to each connection is a reasonable, abstract model of this reduced connectedness. The mechanisms underlying biological spine loss are a subject of debate (Dorostkar et al., 2015; Fiala et al., 2002), making it difficult to comment on the biological relevance of weight decay. It does seem that any mechanism which results in random pruning of spines or decreased probability of spine formation would, statistically, affect the likelihood of strong connections (involving many synapses) more than weak connections—an effect similar to that of weight decay, which degrades the strongest connections fastest. Similarly, mitochondrial dysfunction or otherwise altered energy dynamics within a neuron would seem to affect the strongest synapses with the highest metabolic cost. But until more is known, weight decay remains attractive mostly because it is reasonable and convenient. Moreover, simulating spine loss by deleting random connections from the network—which is also reasonable and has been used to model other disorders that involve spine loss (Lanillos et al., 2020; Tuladhar et al., 2021)—does not seem to produce the same range of depression-like behaviors in our experiments (see Supplementary material). This may be because the loss of spines is part of a bigger picture of impaired neural health and communication, and weight decay better models this net degradation of communication channels.

We have shown the weight decay model already produces some interesting results. But even more insight may be available if future work can devise a more nuanced and biologically accurate way to simulate spine loss. For example, we have applied weight decay throughout the entire network, but in reality it seems spine density changes depend on brain region (Qiao et al., 2016; Runge et al., 2020) and types of play experience (Bell et al., 2010; Stark et al., 2023). Thus our computational framework is primarily a model of the reduced

spine density observed in the prefrontal cortex and hippocampus during MDD. That is, our model highlights this particular aspect of the complex disorder—and produces a surprising range of depression-like results by doing so. Making the simulations more biologically accurate in ways that account for varied effects in different brain regions would require significant innovation in the artificial neural network approach, but may create a new range of hypotheses.

The involvement of microglia on MDD pathology is another important aspect of MDD falling outside the scope of our model. Microglia are cells that serve an immune function in the central nervous system. MDD seems to involve microglia function that is altered in a variety of ways, and microglia likely have a significant role in MDD pathology (Snijders et al., 2021; Wang et al., 2022). Our model focuses on the view of MDD as a dysfunction of neuroplasticity, while partially accounting for the dopamine system dysfunction that was part of the older monoamine hypothesis. But seeing MDD as a microglial disease is another important view (Yirmiya et al., 2015)—though it falls outside the scope of our present model. The two views are certainly not orthogonal: microglia regulate synaptic plasticity, synaptic refinement and pruning, and formation of neural networks (Wang et al., 2022)—all of which are taken for granted in our computational model. That is, our model simulates spine loss but abstracts away the specific role that microglia may play in this process. Future computational work should strive for a more detailed and comprehensive account.

The spine loss modeled here is one of many changes that occur across various brain regions in MDD. Some of these changes seem conceptually or effectively related to spine loss [for example, along with loss of spines in the hippocampus, reduced neurogenesis in the hippocampus has also been observed (Berger et al., 2020)] while others are not. Here we hypothesized that among all the changes, spine loss is particularly important, and our modeling shows that it can account for an interesting range of phenotypes found in MDD.

All models exist on a scale from high abstraction to high detail: abstracting away complexity illustrates general "big picture" principles, which is appropriate for a first-of-its-kind study such as this one. We hope future work (our own and others') will build more detailed models that complement (or indeed replace) this one.

Questions of multifinality

Spine or synapse loss is a feature of depression, but also of disorders such as Parkinson's disease (Gcwensa et al., 2021), Alzheimer's disease (where spine loss occurs in clusters and is paired with neuronal cell death (Goel et al., 2022; Mijalkov et al., 2021)), and schizophrenia [where abnormal pruning leaves an excitation-inhibition imbalance (Liu et al., 2021)]. Each of these disorders seems to involve a unique pattern of loss, which should be expected to create unique symptoms.

Interestingly, while we have used weight decay to simulate spine loss, and observed depression-like effects, other researchers have used related approaches to simulate Alzheimer's disease (Tuladhar et al., 2021), and schizophrenia and autism spectrum disorder (Lanillos et al., 2020). Why do subtle variations on simulated spine loss seem to account for so many aspects of these varied disorders? Hayato Idei has begun to explore this question of "multifinality" through neurorobotics (Idei and Yamashita, 2024), but this is a question the emerging field of

RL *in-silico* models must address. Certainly altered synapse density is a feature of many different disorders—are our various observations from *in-silico* models suggesting that altered spine density is a *key* feature across multiple pathologies? Or do *in silico* models simply lack the power to tease out subtle differences between related disorders?

Model validation

All disease models, whether mathematical, schematic, computational, or animal, must be validated to identify how and how effectively they represent the disease. Since models are metaphors or abstractions, the validation process never shows them to be exact matches of the disease, but rather identifies the particular features of the disease that the model represents. Models range from high-level abstractions like our RL model, to detailed biological analogs like an animal model. Animal models could be validated by three main criteria: (1) similarity of symptoms, (2) similarity of etiology or causative factors, and (3) similar outcomes of pharmacological interventions. *In-silico* models such as ours exist at a higher level of abstraction and serve a more illustrative purpose. They are typically evaluated by how well they reflect known biological mechanisms and produce expected symptoms, and through comparison to other existing in-silico models. Our model reflects the known biological mechanism of spine loss (which is not to say that other known mechanisms are not important, only that this is the focus of our particular model) and produces a range of expected depression-like symptoms. It compares favorably with other models, in the sense that it better produces the expected symptoms. While this is sufficient to make the model an interesting contribution, future work could perform a deeper level of validation by fitting similar models to behavioral data, or testing the various hypotheses and falsifiable predictions made by the model, as described above.

Methods

Computational model

Our deep reinforcement learning agent is based on temporal difference (TD) learning, which supposes that an agent inhabits some state *s*, and can select an action to perform from some set of actions *A*. If executing that action leads the agent to a new state *s'*, and triggers some reward *r*, then the value of that particular experience can be formulated as in Equation 1:

$$r + \gamma \cdot \max_{b \in A}\left[V\left(s',b\right)\right] \qquad 1$$

where γ is a discount factor (between 0 and 1) that discounts the value of future rewards relative to immediate ones. Note that *r* can be zero (no reward) or negative (a punishment). The value *in general* of executing action *a* from state *s* is then the average or expected value of all these individual experiences, as expressed in Equation 2:

$$V\left(s,a\right) = \mathrm{E}\left[r + \gamma \cdot \max_{b \in A}\left[V\left(s',b\right)\right]\right] \qquad 2$$

The learning process consists of updating the value estimates V after each experience in the world. A difference δ is computed between the experienced value and the current value estimate, and then used to adjust the estimate according to a learning rate α:

$$\delta = r + \gamma \cdot \max_{b \in A}\left[V\left(s',b\right)\right] - V\left(s,a\right) \qquad 3$$

$$V\left(s,a\right)_{new} = V\left(s,a\right)_{old} + \alpha \cdot \delta \qquad 4$$

The difference δ is the reward-prediction-error that is thought to be signaled by dopamine neurons (Montague et al., 1996; Schultz et al., 1997). For more information on TD learning see Sutton and Barto (2018). Here we take a Deep Reinforcement Learning approach similar to that of Mnih et al. (2015), in which an artificial neural network creates the value estimates $V(s,a)$. This network accepts an input of 100 binary values (corresponding to the agent's 5×5 visual field, times 4 object types that can exist at each location within the field). The network has a hidden layer of 10 neurons with tanh activation functions, and an output layer of 3 linear neurons whose outputs represent the values of turning left and right and moving forward.

The learning process now consists of tuning network connection weights such that the value estimates become increasingly accurate. For each experience, δ is computed per Equation 3, the gradient of δ^2 is calculated with respect to each network weight w_i, and the weight is adjusted in the direction that will minimize δ:

$$w_{i_{new}} = w_{i_{old}} - \alpha \frac{\partial \delta^2}{\partial w_{i_{old}}} \qquad 5$$

Note the analogy between Equations 4 and 5. The weight decay effect is created by adding a sum of square weights, $\frac{1}{2}\lambda \mathbf{w}^2$, to δ^2 before computing the gradient. Equation 6 then becomes:

$$w_{i_{new}} = w_{i_{old}} - \alpha \frac{\partial \delta^2}{\partial w_{i_{old}}} - \alpha\lambda w_{i_{old}} \qquad 6$$

where λ controls how quickly each weight decays toward zero. A small amount of weight decay is often used with artificial neural networks to improve their performance and generalization (Andriushchenko et al., 2023), but here we use a more extreme weight decay setting to simulate significant spine loss.

See Supplementary material for all parameter settings.

Alternative model implementations

Our "healthy" agent was identical to the agent with simulated spine loss, except that the healthy agent used a weight decay setting of $\lambda=0$ (no weight decay). The variety of alternative computational MDD models shown in Table 1 were implemented as follows:

Reduced reward-prediction-error signal: For this model, the reward prediction error signal δ was multiplied by a scale factor of 0.1 to simulate reduced dopamine signaling. Note that this scaling in the TD learning model is mathematically identical to using a smaller learning rate α.

Faster learning from negative experiences: This model multiplied all negative rewards by 2 and divided all positive rewards by 2 before doing network weight updates. Thus, negative rewards had a greater effect on those updates.

Higher discount factor: This model used a discount factor of $\gamma = 0.5$, rather than the $\gamma = 0.9$ used by other models. Per Equation 1, this increases the discounting of future rewards.
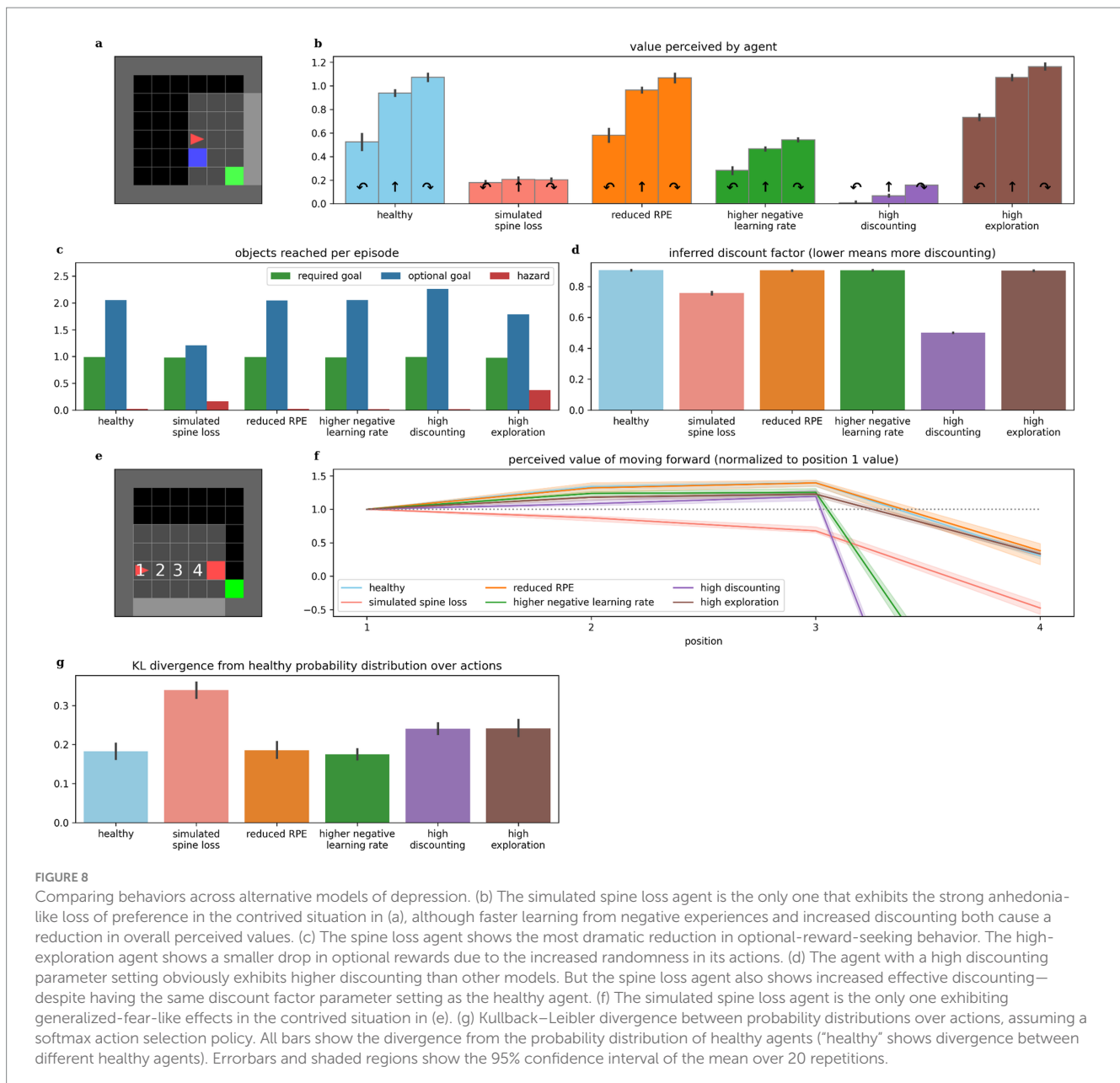
Increased exploration: Studies that use reinforcement learning to model animal behavior have often assumed a softmax action selection policy (Cinotti et al., 2019; Dongelmans et al., 2021; Ohta et al., 2021). Under this policy, the agent is assumed to select actions stochastically; actions with higher perceived value have a higher probability of selection:

$$P\left(a|s\right) = \frac{e^{v\left(s,a\right)/\tau}}{\sum_{b \in A} e^{V\left(s,b\right)/\tau}} \qquad 7$$

Here τ is a "temperature" parameter that sets the strength of the agent's preference for exploiting the action with highest value, over exploring the other actions (for example, for high "temperature" values, the agent's choice of action will be closer to random). Our "increased exploration" model uses this softmax action selection with $\tau = 1$. This produces many more exploratory actions than the other models, which use an "ε-greedy" strategy of selecting the highest-value action with probability 1-ε, and a random (exploratory) action otherwise.

Comparison with alternative depression models

Comparisons between the spine loss model and the other depression models are illustrated in Figure 8, and were used to construct Table 1. Models were considered to demonstrate anhedonia if there was no significant difference between the perceived values for turning right and moving forward in the experiment from Figure 2d. Models were considered to exhibit generalized avoidance/fear if the perceived value of moving toward the hazard decreased prematurely, as in Figure 3d. The spine loss model was the only model to exhibit anhedonia and generalized avoidance by these criteria, although the model with high discounting does exhibit lower perceived values in general (due to the extra discounting during evaluation of Equation 3). Inferred discount factors were measured as in Figures 3a,b. To evaluate relative differences in exploration rates, probability distributions were generated over actions (using Equation 7). The distribution for the healthy agent was compared to those for other models using Kullback–Leibler divergence, for a number of randomly-generated world states. This approach uses the healthy agent's probability distribution as a baseline, and measures deviation from that distribution in the other agents.

FIGURE 8
Comparing behaviors across alternative models of depression. (b) The simulated spine loss agent is the only one that exhibits the strong anhedonia-like loss of preference in the contrived situation in (a), although faster learning from negative experiences and increased discounting both cause a reduction in overall perceived values. (c) The spine loss agent shows the most dramatic reduction in optional-reward-seeking behavior. The high-exploration agent shows a smaller drop in optional rewards due to the increased randomness in its actions. (d) The agent with a high discounting parameter setting obviously exhibits higher discounting than other models. But the spine loss agent also shows increased effective discounting—despite having the same discount factor parameter setting as the healthy agent. (f) The simulated spine loss agent is the only one exhibiting generalized-fear-like effects in the contrived situation in (e). (g) Kullback–Leibler divergence between probability distributions over actions, assuming a softmax action selection policy. All bars show the divergence from the probability distribution of healthy agents ("healthy" shows divergence between different healthy agents). Errorbars and shaded regions show the 95% confidence interval of the mean over 20 repetitions.

Significance statement

Simulating dendritic spine loss in a deep reinforcement learning agent causes the agent to exhibit a surprising range of depression-like behaviors. Simulating spine restoration allows rewarding behavior to be re-learned. This computational model sees Major Depressive Disorder as a reversible loss of brain capacity, providing some insights on pathology and treatment.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://github.com/echalmers/blue_ai/releases/tag/v1.0.0.

Author contributions

EC: Writing – review & editing, Writing – original draft, Supervision, Software, Methodology, Investigation, Funding acquisition, Conceptualization. SD: Writing – review & editing, Writing – original draft, Investigation, Conceptualization. XA-H: Writing – review & editing, Writing – original draft, Investigation, Conceptualization. DD: Writing – review & editing, Conceptualization. AG: Writing – review & editing, Supervision, Methodology, Conceptualization. RM: Writing – review & editing, Supervision, Methodology, Conceptualization.

Funding

Acknowledgments

Conflict of interest

The 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.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2024.1466364/full#supplementary-material

References

Aleksandrova, L. R., and Phillips, A. G. (2021). Neuroplasticity as a convergent mechanism of ketamine and classical psychedelics. *Trends Pharmacol. Sci.* 42, 929–942. doi: 10.1016/j.tips.2021.08.003

American Psychiatric Association (2022). Diagnostic and statistical manual of mental disorders. *DSM-5-TR.* Edn. Washington, DC: American Psychiatric Association Publishing.

Amlung, M., Marsden, E., Holshausen, K., Morris, V., Patel, H., Vedelago, L., et al. (2019). Delay discounting as a Transdiagnostic process in psychiatric disorders: a meta-analysis. *JAMA Psychiatry* 76, 1176–1186. doi: 10.1001/jamapsychiatry.2019.2102

Andriushchenko, M., D'Angelo, F., Varre, A., and Flammarion, N., (2023). Why do we need weight decay in modern deep learning? doi: 10.48550/arXiv.2310.04415

Bakic, J., Pourtois, G., Jepma, M., Duprat, R., De Raedt, R., and Baeken, C. (2017). Spared internal but impaired external reward prediction error signals in major depressive disorder during reinforcement learning. *Depress. Anxiety* 34, 89–96. doi: 10.1002/da.22576

Bell, H. C., Pellis, S. M., and Kolb, B. (2010). Juvenile peer play experience and the development of the orbitofrontal and medial prefrontal cortices. *Behav. Brain Res.* 207, 7–13. doi: 10.1016/j.bbr.2009.09.029

Belujon, P., and Grace, A. A. (2017). Dopamine system dysregulation in major depressive disorders. *Int. J. Neuropsychopharmacol.* 20, 1036–1046. doi: 10.1093/ijnp/pyx056

Berger, T., Lee, H., Young, A. H., Aarsland, D., and Thuret, S. (2020). Adult hippocampal neurogenesis in major depressive disorder and Alzheimer's disease. *Trends Mol. Med.* 26, 803–818. doi: 10.1016/j.molmed.2020.03.010

Berman, R. M., Cappiello, A., Anand, A., Oren, D. A., Heninger, G. R., Charney, D. S., et al. (2000). Antidepressant effects of ketamine in depressed patients. *Biol. Psychiatry* 47, 351–354. doi: 10.1016/s0006-3223(99)00230-9

Bessa, J. M., Morais, M., Marques, F., Pinto, L., Palha, J. A., and Almeida, O. F. X., et al. (2013). Stress-induced anhedonia is associated with hypertrophy of medium spiny neurons of the nucleus accumbens. Transl. Psychiatry 3,:e266. doi: 10.1038/tp.2013.39

Blanco, N. J., Otto, A. R., Maddox, W. T., Beevers, C. G., and Love, B. C. (2013). The influence of depression symptoms on exploratory decision-making. *Cognition* 129, 563–568. doi: 10.1016/j.cognition.2013.08.018

Botvinick, M., Wang, J. X., Dabney, W., Miller, K. J., and Kurth-Nelson, Z. (2020). Deep reinforcement learning and its neuroscientific implications. *Neuron* 107, 603–616. doi: 10.1016/j.neuron.2020.06.014

Broekens, J., Jacobs, E., and Jonker, C. M. (2015). A reinforcement learning model of joy, distress, hope and fear. *Connect. Sci.* 27, 215–233. doi: 10.1080/09540091.2015.1031081

Brolsma, S. C. A., Vrijsen, J. N., Vassena, E., Kandroodi, M. R., Bergman, M. A., Eijndhoven, P. F.Van, et al. (2022). Challenging the negative learning bias hypothesis of depression: reversal learning in a naturalistic psychiatric sample. *Psychol. Med.* 52, 303–313. doi: 10.1017/S0033291720001956

Browne, C. A., and Lucki, I. (2013). Antidepressant effects of ketamine: mechanisms underlying fast-acting novel antidepressants. *Front. Pharmacol.* 4:161. doi: 10.3389/fphar.2013.00161

Calder, A. E., and Hasler, G. (2023). Towards an understanding of psychedelic-induced neuroplasticity. *Neuropsychopharmacology* 48, 104–112. doi: 10.1038/s41386-022-01389-z

Castrén, E., and Monteggia, L. M. (2021). Brain-derived neurotrophic factor signaling in depression and antidepressant action. *Biol. Psychiatry, Ronald Duman's Legacy* 90, 128–136. doi: 10.1016/j.biopsych.2021.05.008

Chung, D., Kadlec, K., Aimone, J. A., McCurry, K., King-Casas, B., and Chiu, P. H. (2017). Valuation in major depression is intact and stable in a non-learning environment. *Sci. Rep.* 7:44374. doi: 10.1038/srep44374

Cinotti, F., Fresno, V., Aklil, N., Coutureau, E., Girard, B., Marchand, A. R., et al. (2019). Dopamine blockade impairs the exploration-exploitation trade-off in rats. *Sci. Rep.* 9:6770. doi: 10.1038/s41598-019-43245-z

Daw, N. (2012). Model-based reinforcement learning as cognitive search: Neurocomputational theories. MIT Press.

Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711. doi: 10.1038/nn1560

Dongelmans, M., Durand-de Cuttoli, R., Nguyen, C., Come, M., Duranté, E. K., Lemoine, D., et al. (2021). Chronic nicotine increases midbrain dopamine neuron activity and biases individual strategies towards reduced exploration in mice. *Nat. Commun.* 12:6945. doi: 10.1038/s41467-021-27268-7

Dorostkar, M. M., Zou, C., Blazquez-Llorca, L., and Herms, J. (2015). Analyzing dendritic spine pathology in Alzheimer's disease: problems and opportunities. *Acta Neuropathol. (Berl.)* 130, 1–19. doi: 10.1007/s00401-015-1449-5

dos Santos, R. G., Osório, F. L., Crippa, J. A. S., Riba, J., Zuardi, A. W., and Hallak, J. E. C. (2016). Antidepressive, anxiolytic, and antiaddictive effects of ayahuasca, psilocybin and lysergic acid diethylamide (LSD): a systematic review of clinical trials published in the last 25 years. *Ther. Adv. Psychopharmacol.* 6, 193–213. doi: 10.1177/2045125316638008

Duman, R. S., Aghajanian, G. K., Sanacora, G., and Krystal, J. H. (2016). Synaptic plasticity and depression: new insights from stress and rapid-acting antidepressants. *Nat. Med.* 22, 238–249. doi: 10.1038/nm.4050

Edinoff, A. N., Akuly, H. A., Hanna, T. A., Ochoa, C. O., Patti, S. J., Ghaffar, Y. A., et al. (2021). Selective serotonin reuptake inhibitors and adverse effects: a narrative review. *Neurol. Int.* 13, 387–401. doi: 10.3390/neurolint13030038

Eshel, N., and Roiser, J. P. (2010). Reward and punishment processing in depression. *Biol. Psychiatry* 68, 118–124. doi: 10.1016/j.biopsych.2010.01.027

Fiala, J. C., Spacek, J., and Harris, K. M. (2002). Dendritic spine pathology: cause or consequence of neurological disorders? *Brain Res. Rev.* 39, 29–54. doi: 10.1016/S0165-0173(02)00158-3

Gcwensa, N. Z., Russell, D. L., Cowell, R. M., and Volpicelli-Daley, L. A. (2021). Molecular mechanisms underlying synaptic and axon degeneration in Parkinson's disease. *Front. Cell. Neurosci.* 15:626128. doi: 10.3389/fncel.2021.626128

Goel, P., Chakrabarti, S., Goel, K., Bhutani, K., Chopra, T., and Bali, S. (2022). Neuronal cell death mechanisms in Alzheimer's disease: an insight. *Front. Mol. Neurosci.* 15:937133. doi: 10.3389/fnmol.2022.937133

Gruber, A. J., Calhoon, G. G., Shusterman, I., Schoenbaum, G., Roesch, M. R., and O'Donnell, P. (2010). More is less: a disinhibited prefrontal cortex impairs cognitive flexibility. *J. Neurosci.* 30, 17102–17110. doi: 10.1523/JNEUROSCI.4623-10.2010

Hallford, D. J., Barry, T. J., Austin, D. W., Raes, F., Takano, K., and Klein, B. (2020). Impairments in episodic future thinking for positive events and anticipatory pleasure in major depression. *J. Affect. Disord.* 260, 536–543. doi: 10.1016/j.jad.2019.09.039

Hanlon, F. M., and Sutherland, R. J. (2000). Changes in adult brain and behavior caused by neonatal limbic damage: implications for the etiology of schizophrenia. *Behav. Brain Res.* 107, 71–83. doi: 10.1016/s0166-4328(99)00114-x

Harmer, C. J., Goodwin, G. M., and Cowen, P. J. (2009a). Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. *Br. J. Psychiatry* 195, 102–108. doi: 10.1192/bjp.bp.108.051193

Harmer, C. J., O'Sullivan, U., Favaron, E., Massey-Chase, R., Ayres, R., Reinecke, A., et al. (2009b). Effect of acute antidepressant administration on negative affective Bias in depressed patients. *Am. J. Psychiatry* 166, 1178–1184. doi: 10.1176/appi.ajp.2009.09020149

Hawkins, J., and Ahmad, S. (2016). Why neurons have thousands of synapses, a theory of sequence memory in neocortex. *Front. Neural Circuits* 10:174222. doi: 10.3389/fncir.2016.00023

Helm, K., Viol, K., Weiger, T. M., Tass, P. A., Grefkes, C., Del Monte, D., et al. (2018). Neuronal connectivity in major depressive disorder: a systematic review. *Neuropsychiatr. Dis. Treat.* 14, 2715–2737. doi: 10.2147/NDT.S170989

Herzallah, M., Moustafa, A., Natsheh, J., Abdellatif, S., Taha, M., Tayem, Y., et al. (2013). Learning from negative feedback in patients with major depressive disorder is attenuated by SSRI antidepressants. *Front. Integr. Neurosci.* 7:7. doi: 10.3389/fnint.2013.00067

Hiratani, N., and Fukai, T. (2018). Redundancy in synaptic connections enables neurons to learn optimally. *Proc. Natl. Acad. Sci.* 115, E6871–E6879. doi: 10.1073/pnas.1803274115

Holmes, S. E., Scheinost, D., Finnema, S. J., Naganawa, M., Davis, M. T., DellaGioia, N., et al. (2019). Lower synaptic density is associated with depression severity and network alterations. *Nat. Commun.* 10:1529. doi: 10.1038/s41467-019-09562-7

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., and Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biol. Mood Anxiety Disord.* 3:12. doi: 10.1186/2045-5380-3-12

Idei, H., and Yamashita, Y. (2024). Elucidating multifinal and equifinal pathways to developmental disorders by constructing real-world neurorobotic models. *Neural Netw.* 169, 57–74. doi: 10.1016/j.neunet.2023.10.005

Imhoff, S., Harris, M., Weiser, J., and Reynolds, B. (2014). Delay discounting by depressed and non-depressed adolescent smokers and non-smokers. *Drug Alcohol Depend.* 135, 152–155. doi: 10.1016/j.drugalcdep.2013.11.014

Ironside, M., Amemori, K., McGrath, C. L., Pedersen, M. L., Kang, M. S., Amemori, S., et al. (2020). Approach-avoidance conflict in major depressive disorder: congruent neural findings in humans and nonhuman Primates. *Biol. Psychiatry* 87, 399–408. doi: 10.1016/j.biopsych.2019.08.022

Jauhar, S., Arnone, D., Baldwin, D. S., Bloomfield, M., Browning, M., Cleare, A. J., et al. (2023). A leaky umbrella has little value: evidence clearly indicates the serotonin system is implicated in depression. *Mol. Psychiatry* 28, 3149–3152. doi: 10.1038/s41380-023-02095-y

Johansen, A., Armand, S., Plavén-Sigray, P., Nasser, A., Ozenne, B., Petersen, I. N., et al. (2023). Effects of escitalopram on synaptic density in the healthy human brain: a randomized controlled trial. *Mol. Psychiatry* 28, 4272–4279. doi: 10.1038/s41380-023-02285-8

Kozachkov, L., Tauber, J., Lundqvist, M., Brincat, S. L., Slotine, J.-J., and Miller, E. K. (2022). Robust and brain-like working memory through short-term synaptic plasticity. *PLoS Comput. Biol.* 18:e1010776. doi: 10.1371/journal.pcbi.1010776

Krystal, J. H., Abdallah, C. G., Sanacora, G., Charney, D. S., and Duman, R. S. (2019). Ketamine: a paradigm shift for depression research and treatment. *Neuron* 101, 774–778. doi: 10.1016/j.neuron.2019.02.005

Krystal, J. H., Kaye, A. P., Jefferson, S., Girgenti, M. J., Wilkinson, S. T., Sanacora, G., et al. (2023). Ketamine and the neurobiology of depression: toward next-generation rapid-acting antidepressant treatments. *Proc. Natl. Acad. Sci.* 120:e2305772120. doi: 10.1073/pnas.2305772120

Kunisato, Y., Okamoto, Y., Ueda, K., Onoda, K., Okada, G., Yoshimura, S., et al. (2012). Effects of depression on reward-based decision making and variability of action in probabilistic learning. *J. Behav. Ther. Exp. Psychiatry* 43, 1088–1094. doi: 10.1016/j.jbtep.2012.05.007

Lanillos, P., Oliva, D., Philippsen, A., Yamashita, Y., Nagai, Y., and Cheng, G. (2020). A review on neural network models of schizophrenia and autism spectrum disorder. *Neural Netw.* 122, 338–363. doi: 10.1016/j.neunet.2019.10.014

Lee, E. E., Torous, J., De Choudhury, M., Depp, C. A., Graham, S. A., Kim, H.-C., et al. (2021). Artificial intelligence for mental health care: clinical applications, barriers, facilitators, and artificial wisdom. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 6, 856–864. doi: 10.1016/j.bpsc.2021.02.001

Lempert, K. M., and Pizzagalli, D. A. (2010). Delay discounting and future-directed thinking in anhedonic individuals. *J. Behav. Ther. Exp. Psychiatry* 41, 258–264. doi: 10.1016/j.jbtep.2010.02.003

Li, J.-Z., Gui, D.-Y., Feng, C.-L., Wang, W.-Z., Du, B.-Q., Gan, T., et al. (2012). Victims' time discounting 2.5 years after the Wenchuan earthquake: An ERP study. *PLoS One* 7:e40316. doi: 10.1371/journal.pone.0040316

Li, N., Lee, B., Liu, R.-J., Banasr, M., Dwyer, J. M., Iwata, M., et al. (2010). mTOR-dependent synapse formation underlies the rapid antidepressant effects of NMDA antagonists. *Science* 329, 959–964. doi: 10.1126/science.1190287

Li, N., Liu, R.-J., Dwyer, J. M., Banasr, M., Lee, B., Son, H., et al. (2011). Glutamate N-methyl-D-aspartate receptor antagonists rapidly reverse behavioral and synaptic deficits caused by chronic stress exposure. *Biol. Psychiatry* 69, 754–761. doi: 10.1016/j.biopsych.2010.12.015

Liu, B., Liu, J., Wang, M., Zhang, Y., and Li, L. (2017). From serotonin to neuroplasticity: evolvement of theories for major depressive disorder. *Front. Cell. Neurosci.* 11:11. doi: 10.3389/fncel.2017.00305

Liu, Y., Ouyang, P., Zheng, Y., Mi, L., Zhao, J., Ning, Y., et al. (2021). A selective review of the excitatory-inhibitory imbalance in schizophrenia: underlying biology, genetics, microcircuits, and symptoms. *Front. Cell Dev. Biol.* 9:664535. doi: 10.3389/fcell.2021.664535

Ly, C., Greb, A. C., Cameron, L. P., Wong, J. M., Barragan, E. V., Wilson, P. C., et al. (2018). Psychedelics promote structural and functional neural plasticity. *Cell Rep.* 23, 3170–3182. doi: 10.1016/j.celrep.2018.05.022

Maddox, W. T., Gorlick, M. A., Worthy, D. A., and Beevers, C. G. (2012). Depressive symptoms enhance loss-minimization, but attenuate gain-maximization in history-dependent decision-making. *Cognition* 125, 118–124. doi: 10.1016/j.cognition.2012.06.011

Maia, T. V., and Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nat. Neurosci.* 14, 154–162. doi: 10.1038/nn.2723

Mayford, M., Siegelbaum, S. A., and Kandel, E. R. (2012). Synapses and memory storage. *Cold Spring Harb. Perspect. Biol.* 4:a005751. doi: 10.1101/cshperspect.a005751

Mehrotra, D., and Dubé, L. (2023). Accounting for multiscale processing in adaptive real-world decision-making via the hippocampus. *Front. Neurosci.* 17:17. doi: 10.3389/fnins.2023.1200842

Mijalkov, M., Volpe, G., Fernaud-Espinosa, I., DeFelipe, J., Pereira, J. B., and Merino-Serrais, P. (2021). Dendritic spines are lost in clusters in Alzheimer's disease. *Sci. Rep.* 11:12350. doi: 10.1038/s41598-021-91726-x

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., et al. (2015). Human-level control through deep reinforcement learning. *Nature* 518, 529–533. doi: 10.1038/nature14236

Moda-Sava, R. N., Murdock, M. H., Parekh, P. K., Fetcho, R. N., Huang, B. S., Huynh, T. N., et al. (2019). Sustained rescue of prefrontal circuit dysfunction by antidepressant-induced spine formation. *Science* 364:eaat8078. doi: 10.1126/science.aat8078

Moncrieff, J., Cooper, R. E., Stockmann, T., Amendola, S., Hengartner, M. P., and Horowitz, M. A. (2023). The serotonin theory of depression: a systematic umbrella review of the evidence. *Mol. Psychiatry* 28, 3243–3256. doi: 10.1038/s41380-022-01661-0

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947. doi: 10.1523/JNEUROSCI.16-05-01936.1996

Mukherjee, D., van Geen, C., and Kable, J. (2023). Leveraging decision science to characterize depression. *Curr. Dir. Psychol. Sci.* 32, 462–470. doi: 10.1177/09637214231194962

Murrough, J. W., Iosifescu, D. V., Chang, L. C., Al Jurdi, R. K., Green, C. E., Perez, A. M., et al. (2013). Antidepressant efficacy of ketamine in treatment-resistant major depression: a two-site randomized controlled trial. *Am. J. Psychiatry* 170, 1134–1142. doi: 10.1176/appi.ajp.2013.13030392

National Institute of Mental Health (n.d.) Major depression. Available at: https://www.nimh.nih.gov/health/statistics/major-depression (Accessed October 30, 2024).

Neftci, E. O., and Averbeck, B. B. (2019). Reinforcement learning in artificial and biological systems. *Nat. Mach. Intell.* 1, 133–143. doi: 10.1038/s42256-019-0025-4

Ohta, H., Satori, K., Takarada, Y., Arake, M., Ishizuka, T., Morimoto, Y., et al. (2021). The asymmetric learning rates of murine exploratory behavior in sparse reward environments. *Neural Netw.* 143, 218–229. doi: 10.1016/j.neunet.2021.05.030

Ottenbreit, N. D., Dobson, K. S., and Quigley, L. (2014). An examination of avoidance in major depression in comparison to social anxiety disorder. *Behav. Res. Ther.* 56, 82–90. doi: 10.1016/j.brat.2014.03.005

Pary, R., Scarff, J. R., Jijakli, A., Tobias, C., and Lippmann, S. (2015). A review of psychostimulants for adults with depression. *Fed. Pract.* 32, 30S–37S.

Pattij, T., and Vanderschuren, L. J. M. J. (2008). The neuropharmacology of impulsive behaviour. *Trends Pharmacol. Sci.* 29, 192–199. doi: 10.1016/j.tips.2008.01.002

Pike, A. C., and Robinson, O. J. (2022). Reinforcement learning in patients with mood and anxiety disorders vs control individuals: a systematic review and meta-analysis. *JAMA Psychiatry* 79, 313–322. doi: 10.1001/jamapsychiatry.2022.0051

Pulcu, E., Trotter, P. D., Thomas, E. J., McFarquhar, M., Juhasz, G., Sahakian, B. J., et al. (2014). Temporal discounting in major depressive disorder. *Psychol. Med.* 44, 1825–1834. doi: 10.1017/S0033291713002584

Qiao, H., Li, M.-X., Xu, C., Chen, H.-B., An, S.-C., and Ma, X.-M. (2016). Dendritic spines in depression: what we learned from animal models. *Neural Plast.* 2016, 1–26. doi: 10.1155/2016/8056370

Read, D., and Read, N. L. (2004). Time discounting over the lifespan. *Organ. Behav. Hum. Decis. Process.* 94, 22–32. doi: 10.1016/j.obhdp.2004.01.002

Reinen, J. M., Whitton, A. E., Pizzagalli, D. A., Slifstein, M., Abi-Dargham, A., McGrath, P. J., et al. (2021). Differential reinforcement learning responses to positive and negative information in unmedicated individuals with depression. *Eur. Neuropsychopharmacol.* 53, 89–100. doi: 10.1016/j.euroneuro.2021.08.002

Romeo, B., Karila, L., Martelli, C., and Benyamina, A. (2020). Efficacy of psychedelic treatments on depressive symptoms: a meta-analysis. *J. Psychopharmacol. Oxf. Engl.* 34, 1079–1085. doi: 10.1177/0269881120919957

Rothkirch, M., Tonn, J., Köhler, S., and Sterzer, P. (2017). Neural mechanisms of reinforcement learning in unmedicated patients with major depressive disorder. *Brain* 140, 1147–1157. doi: 10.1093/brain/awx025

Rucker, J. J., Jelen, L. A., Flynn, S., Frowde, K. D., and Young, A. H. (2016). Psychedelics in the treatment of unipolar mood disorders: a systematic review. *J. Psychopharmacol. (Oxf.)* 30, 1220–1229. doi: 10.1177/0269881116679368

Runge, K., Cardoso, C., and De Chevigny, A. (2020). Dendritic spine plasticity: function and mechanisms. *Front. Synaptic Neurosci.* 12:36. doi: 10.3389/fnsyn.2020.00036

Rupprechter, S., Stankevicius, A., Huys, Q. J. M., Steele, J. D., and Seriès, P. (2018). Major depression impairs the use of reward values for decision-making. *Sci. Rep.* 8:13798. doi: 10.1038/s41598-018-31730-w

Rygula, R., Noworyta-Sokolowska, K., Drozd, R., and Kozub, A. (2018). Using rodents to model abnormal sensitivity to feedback in depression. *Neurosci. Biobehav. Rev.* 95, 336–346. doi: 10.1016/j.neubiorev.2018.10.008

Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., et al. (2020). Mastering Atari, go, chess and shogi by planning with a learned model. *Nature* 588, 604–609. doi: 10.1038/s41586-020-03051-4

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593

Smith, A. J., Becker, S., and Kapur, S. (2005). A computational model of the functional role of the ventral-striatal D2 receptor in the expression of previously acquired behaviors. *Neural Comput.* 17, 361–395. doi: 10.1162/0899766053011546

Smoski, M. J., Felder, J., Bizzell, J., Green, S. R., Ernst, M., Lynch, T. R., et al. (2009). fMRI of alterations in reward selection, anticipation, and feedback in major depressive disorder. *J. Affect. Disord.* 118, 69–78. doi: 10.1016/j.jad.2009.01.034

Snijders, G. J. L. J., Sneeboer, M. A. M., Fernández-Andreu, A., Udine, E., Boks, M. P., Ormel, P. R., et al. (2021). Distinct non-inflammatory signature of microglia in post-mortem brain tissue of patients with major depressive disorder. *Mol. Psychiatry* 26, 3336–3349. doi: 10.1038/s41380-020-00896-z

Sojitra, R. B., Lerner, I., Petok, J. R., and Gluck, M. A. (2018). Age affects reinforcement learning through dopamine-based learning imbalance and high decision noise—not through parkinsonian mechanisms. *Neurobiol. Aging* 68, 102–113. doi: 10.1016/j.neurobiolaging.2018.04.006

Speranza, L., Labus, J., Volpicelli, F., Guseva, D., Lacivita, E., Leopoldo, M., et al. (2017). Serotonin 5-HT7 receptor increases the density of dendritic spines and facilitates synaptogenesis in forebrain neurons. *J. Neurochem.* 141, 647–661. doi: 10.1111/jnc.13962

Stark, R. A., Brinkman, B., Gibb, R. L., Iwaniuk, A. N., and Pellis, S. M. (2023). Atypical play experiences in the juvenile period has an impact on the development of the medial prefrontal cortex in both male and female rats. *Behav. Brain Res.* 439:114222. doi: 10.1016/j.bbr.2022.114222

Sutton, R. S., and Barto, A. G. (2018). Reinforcement learning: An introduction. *2nd* Edn. Cambridge, Massachusetts: MIT Press.

Tremblay, L. K., Naranjo, C. A., Cardenas, L., Herrmann, N., and Busto, U. E. (2002). Probing brain reward system function in major depressive disorder: altered response to Dextroamphetamine. *Arch. Gen. Psychiatry* 59, 409–416. doi: 10.1001/archpsyc.59.5.409

Tuladhar, A., Moore, J. A., Ismail, Z., and Forkert, N. D. (2021). Modeling neurodegeneration in silico with deep learning. *Front. Neuroinform.* 15:748370. doi: 10.3389/fninf.2021.748370

Vandendriessche, H., Demmou, A., Bavard, S., Yadak, J., Lemogne, C., Mauras, T., et al. (2023). Contextual influence of reinforcement learning performance of depression: evidence for a negativity bias? *Psychol. Med.* 53, 4696–4706. doi: 10.1017/S0033291722001593

Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dudzik, A., Chung, J., et al. (2019). Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* 575, 350–354. doi: 10.1038/s41586-019-1724-z

Wang, H., He, Y., Sun, Z., Ren, S., Liu, M., Wang, G., et al. (2022). Microglia in depression: an overview of microglia in the pathogenesis and treatment of depression. *J. Neuroinflammation* 19:132. doi: 10.1186/s12974-022-02492-0

Whelan, R., and McHugh, L. A. (2009). Temporal discounting of hypothetical monetary rewards by adolescents, adults, and older adults. *Psychol. Rec.* 59, 247–258. doi: 10.1007/BF03395661

Wong, S. A., Thapa, R., Badenhorst, C. A., Briggs, A. R., Sawada, J. A., and Gruber, A. J. (2017). Opposing effects of acute and chronic d-amphetamine on decision-making in rats. *Neuroscience* 345, 218–228. doi: 10.1016/j.neuroscience.2016.04.021

World Health Organization (Ed.) (1993). The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research. Geneva: World Health Organization.

Yang, C., Hu, Y.-M., Zhou, Z.-Q., Zhang, G.-F., and Yang, J.-J. (2013). Acute administration of ketamine in rats increases hippocampal BDNF and mTOR levels during forced swimming test. *Ups. J. Med. Sci.* 118, 3–8. doi: 10.3109/03009734.2012.724118

Yirmiya, R., Rimmerman, N., and Reshef, R. (2015). Depression as a microglial disease. *Trends Neurosci.* 38, 637–658. doi: 10.1016/j.tins.2015.08.001

Zarate, C. A., Singh, J. B., Carlson, P. J., Brutsche, N. E., Ameli, R., Luckenbaugh, D. A., et al. (2006). A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch. Gen. Psychiatry* 63, 856–864. doi: 10.1001/archpsyc.63.8.856

Zhou, W., Wang, N., Yang, C., Li, X.-M., Zhou, Z.-Q., and Yang, J.-J. (2014). Ketamine-induced antidepressant effects are associated with AMPA receptors-mediated upregulation of mTOR and BDNF in rat hippocampus and prefrontal cortex. *Eur. Psychiatry* 29, 419–423. doi: 10.1016/j.eurpsy.2013.10.005

Zhuang, P.-C., Tan, Z.-N., Jia, Z.-Y., Wang, B., Grady, J. J., and Ma, X.-M. (2019). Treadmill exercise reverses depression model-induced alteration of dendritic spines in the brain areas of mood circuit. *Front. Behav. Neurosci.* 13:93. doi: 10.3389/fnbeh.2019.00093