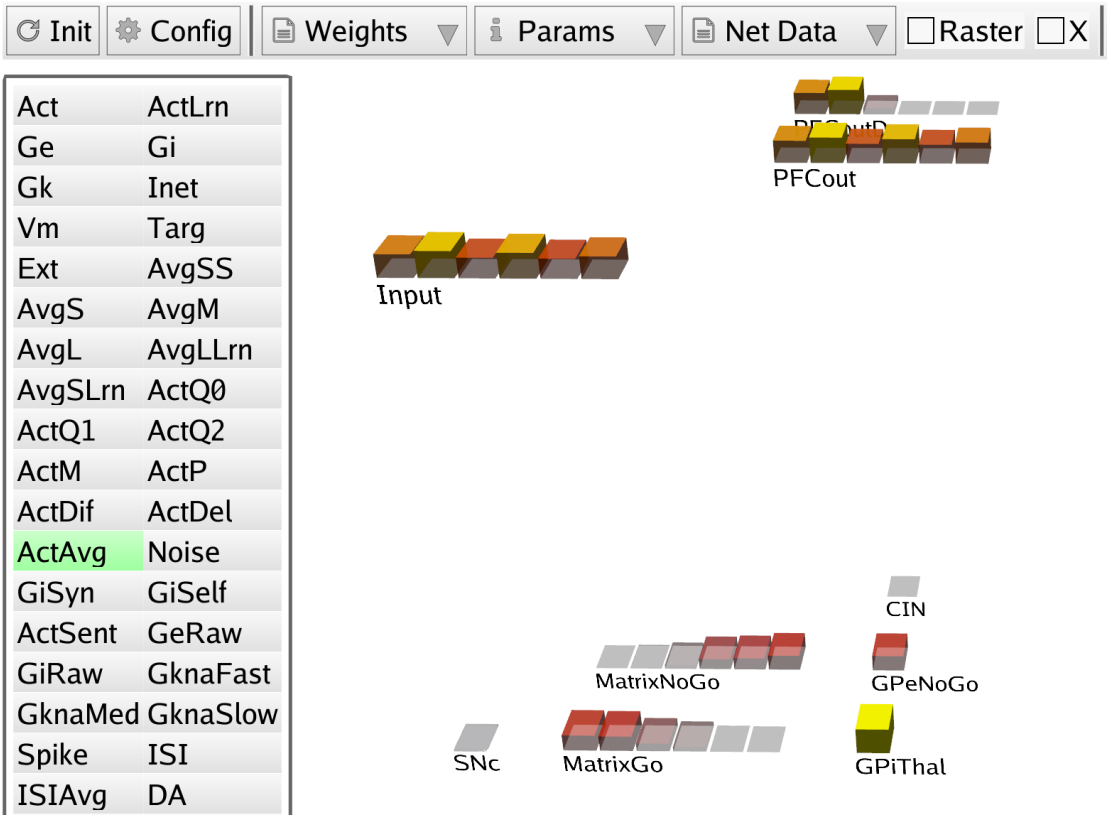
Computational Cognitive Neuroscience – CH7

Zhuo Wang ScM BME Brown ID# 140641091

Question 7.1

The activation patterns of MatrixGo and MatrixNoGo are opposite, PFCout has uniform activation but PFCoutD has only some activation.



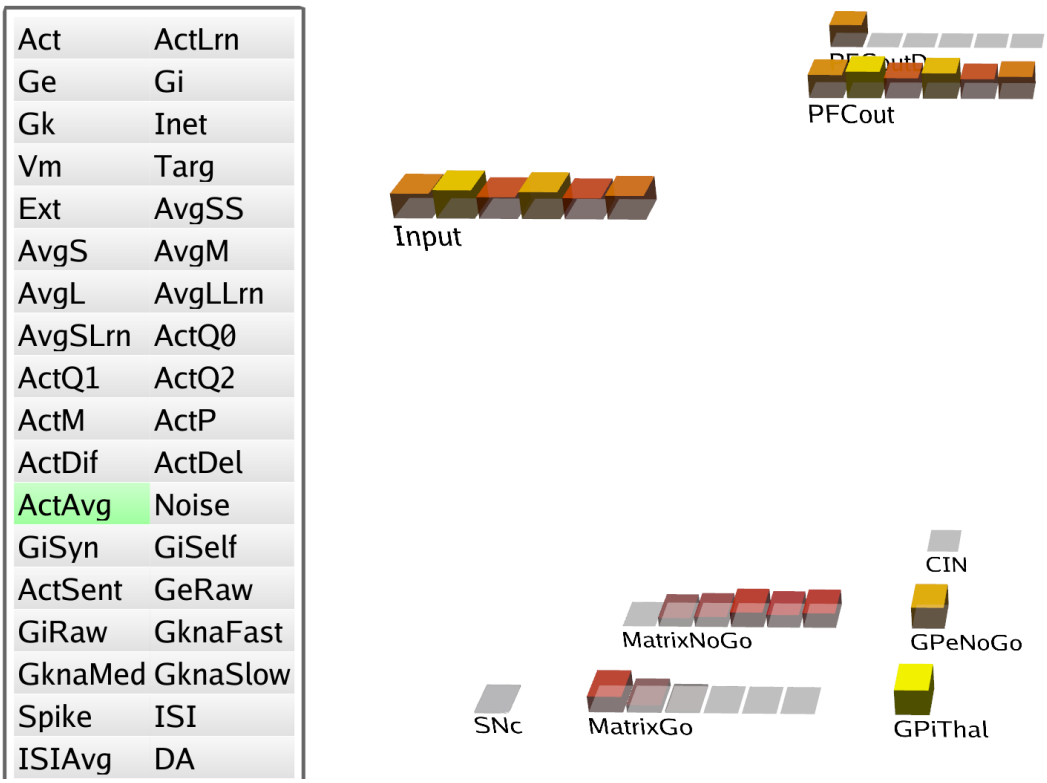
MatrixGo is activated by dopamine bursts during learning. This is because Dopamine bursts, via D1 receptors, cause increased excitability in Go units that are already active due to corticostriatal input. This leads to the reinforcement of actions associated with positive outcomes, as reflected in the increased activation of MatrixGo.

For MatrixNoGo is just the opposite. So in the activation of six inputs they are opposite.

PFCoutD features a unique activation. Deep-layer “output” neurons in the frontal region are represented by PFCoutD. According to some activations, certain actions are selected in accordance with the overall assessment of their results. This layer represents the decision to do a motor action and is triggered when the GPiThal unit exceeds a threshold level, enabling the transition from superficial to deep activation. Therefore, its activation pattern seems to have a positive correlation with the activation pattern of MatrixGo.

Question 7.2

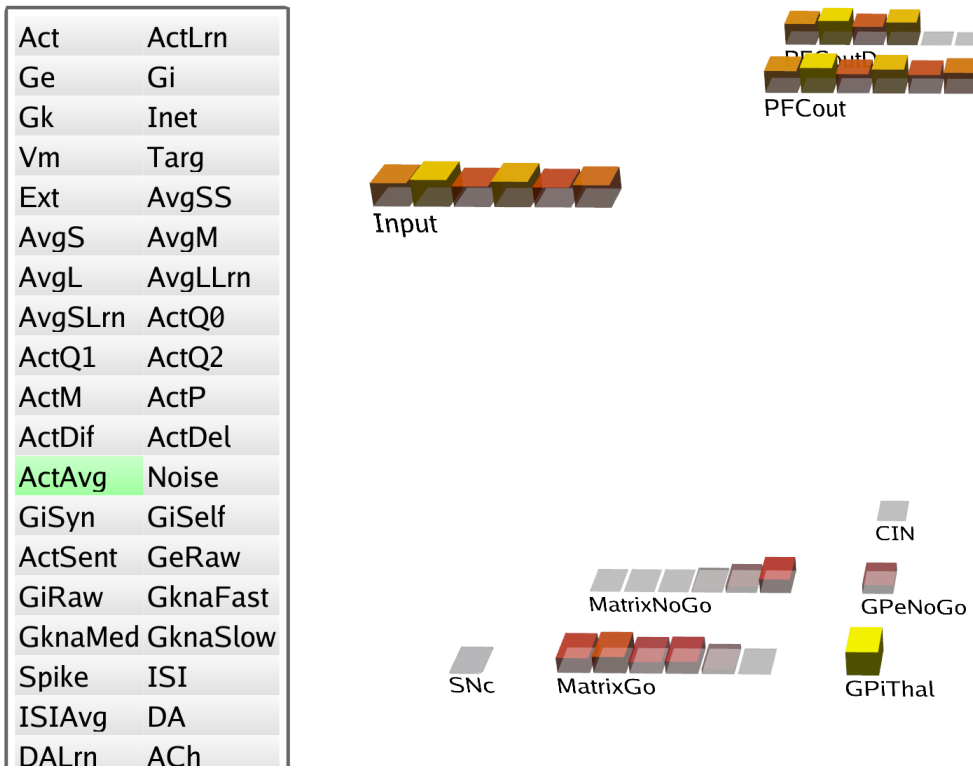
Compared with the “intact” network, the activation patterns of MatrixGo become less, on the contrary, the activation patterns of MatrixNoGo become more, and the activation patterns of PFCoutD become less like MatrixGo.



When PD Patients OFF Meds they were more likely to learn to avoid the B stimulus that was only rewarded 20% of the time. This is consistent with increased activation patterns in MatrixNoGo, suggesting a bias toward dopamine-dip-based learning, as seen in the model with BurstDaGain=0.25. Results of the model show a bias toward NoGo learning is consistent with the observed rise in MatrixNoGo activation. This fits the trend seen in PD patients who are not taking their meds.

Question 7.3

Compared with the “intact” network, the activation patterns of MatrixGo become more, on the contrary, the activation patterns of MatrixNoGo become less, and the activation patterns of PFCoutD become more like MatrixGo.



PD patients who were on meds learned more about selecting the A stimulus, which resulted in an 80% reward rate. As observed in the model with DipDaGain=0.25, this is consistent with enhanced activation patterns in MatrixGo, suggesting a bias toward burst-based learning. According to the model's results, there is a bias toward Go learning and less NoGo learning, as seen by the observed increase in MatrixGo activation and decrease in MatrixNoGo activation. This pattern is comparable to what is seen in PD patients who are on meds.

1

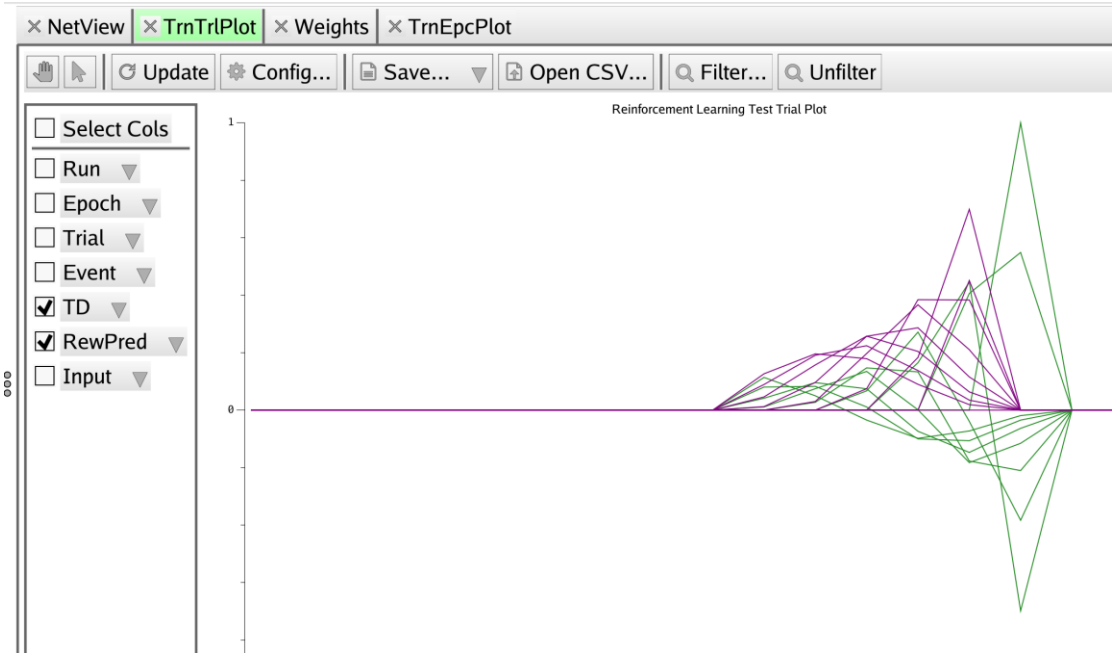
Question 7.5

Based on the learned relationship with the conditioned stimulus CSA, the RewPred layer accurately predicted the reward at this time step prior to deactivating the US at time step 15. Anticipating a reward, the TD unit, which calculates the dopamine-like signal, displayed a positive spike at the anticipated time step.

1 The US was deactivated at time step 15, indicating that the expected reward would not be delivered. This deactivation led to a discrepancy between the expected reward (predicted by the RewPred layer) and the actual reward (deactivated US).

The difference between expected and actual rewards triggered a response from the TD unit, which is in charge of calculating the dopamine-like signal. The expected reward did not materialize as expected, as indicated by the TD delta value turning negative, which represents the reward prediction mistake.

Question 7.6

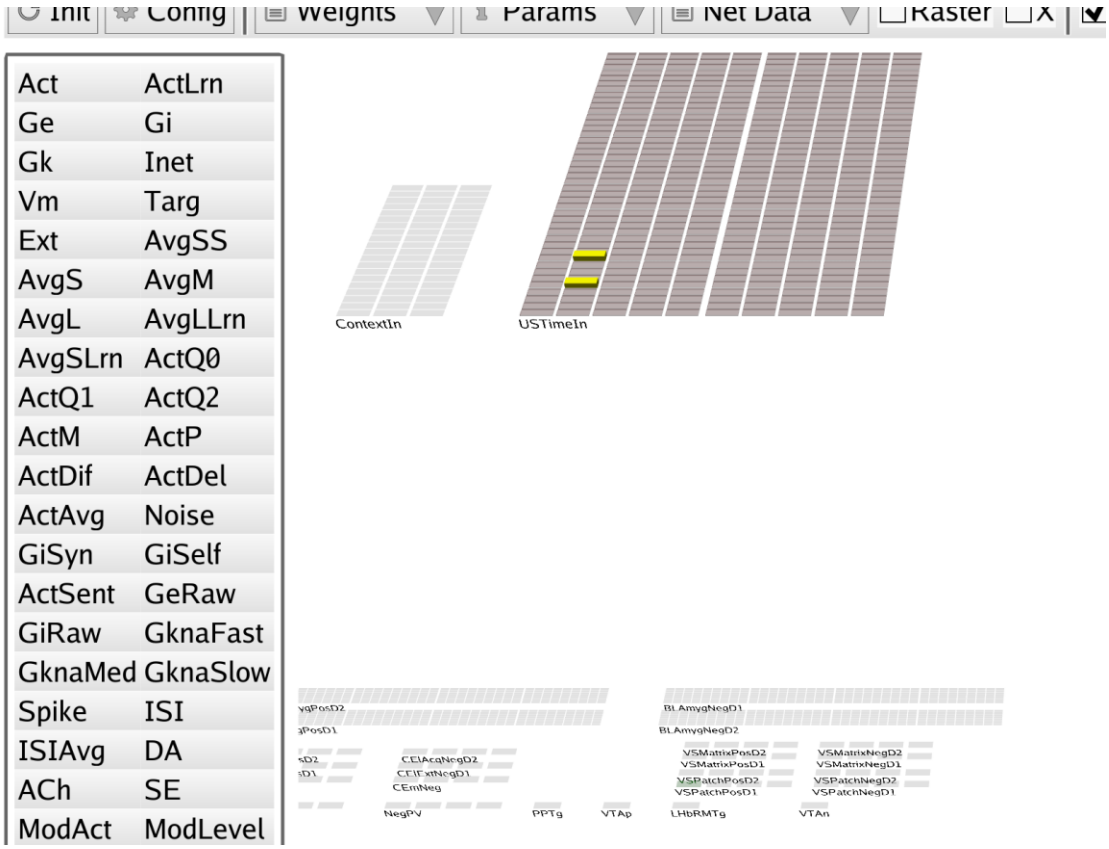


The TD error signals presented in the graph view exhibit a diminish trend over the course of the subsequent many epochs of extinction training. Initially, the TD error signal will be large and negative, indicating a considerable prediction mistake, when 1 the reward (US) is deactivated at the anticipated time step. Over the course of epochs, the TD error signal should diminish as the network goes through extinction training. The network eventually modifies its expectations to align with the altered environment as it recognizes that the CSA is no longer a dependable predictor of the reward, as evidenced by the drop in TD error signals.

This is because the network modifies its expectations in response to the predicted reward not occurring during extinction training, resulting in a decrease in the TD error signal. A lower anticipation of reward in response to this stimulus is indicated by the

weights associated with the CSA decreasing with a negative TD signal. The network adjusts to the new learning environment as extinction training goes on, and the TD error signals get closer to zero.

Question 7.7



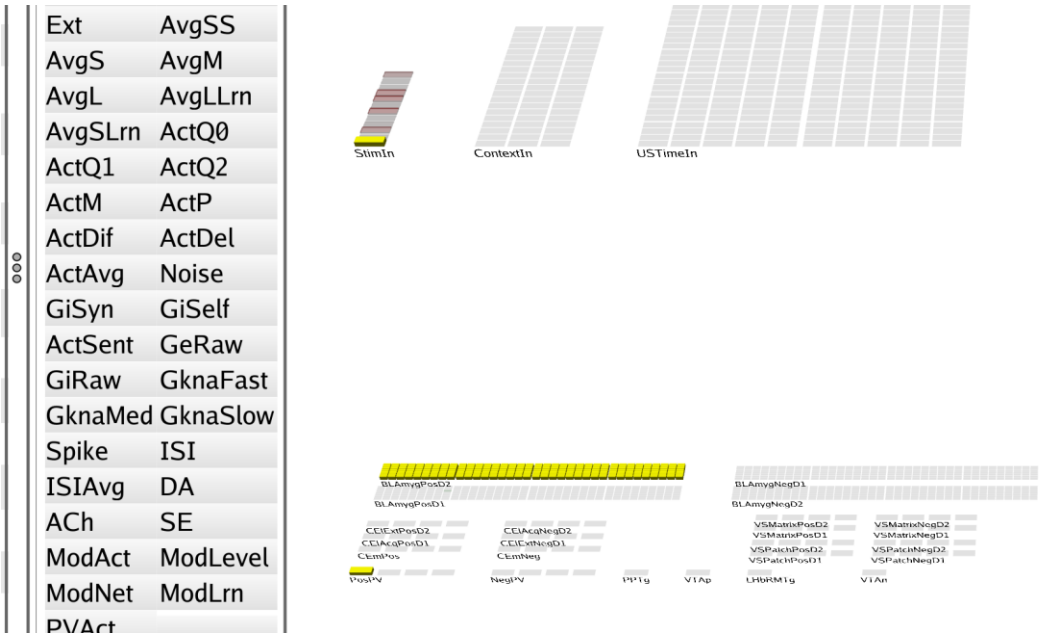
Weights from some units in the USTimeIn layer are sent to the VSPatchPosD1 layer in the PVLV model. Specifically, these weights are from the units in the USTimeIn layer that were active when the US was provided. The mechanism for the network to learn and forecast the timing of reward delivery is provided by the weights from USTimeIn to VSPatchPosD1.

One way of establishing temporal association is by the weights from USTimeIn to VSPatchPosD1. Weights from an active USTimeIn unit to the VSPatchPosD1 layer are enhanced during learning when that unit is active at the moment of reward presentation. Consequently, the enhanced weights cause an increase in activity in the VSPatchPosD1 layer when the CS is delivered in following trials and activates the corresponding USTimeIn unit.

Sending inhibitory signals to the VTAp layer, the VSPatchPosD1 layer becomes increasingly active in anticipation of the reward. This suppression functions as a counterbalance to the usual dopamine surge that follows unexpected benefits. Stated differently, the network's capacity to modify the phasic dopamine response in response

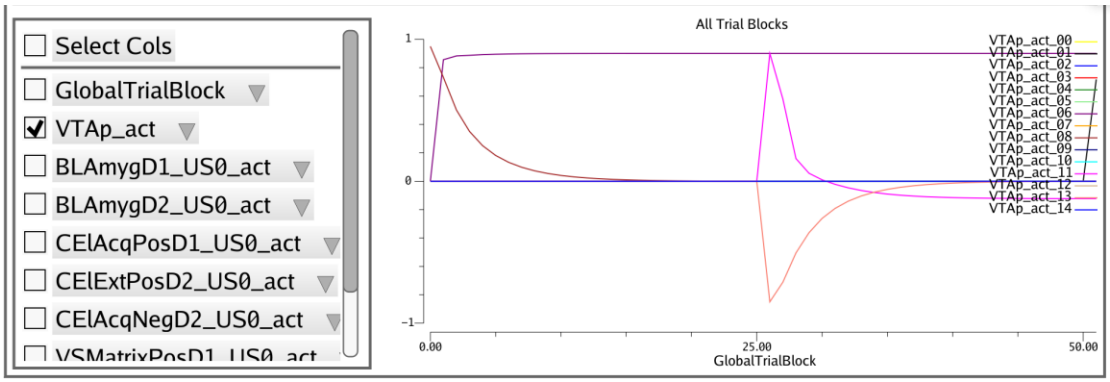
to learnt expectancies is facilitated by the reinforced weights from USTimeIn to VSPatchPosD1, which in turn helps the network anticipate and lessen the burst during the expected reward.

Question 7.8



After extinction training, the units in the BLAmgPosD1 layer still have high weights from StimIn because the goal of extinction learning is to create a new association that prevents the expression of the conditioned response rather than to erase the prior learning. Excitatory weights were formed during acquisition as BLAmgPosD1 learned to correlate the CS (StimIn) with the US. The initial excitatory weights are not entirely eliminated even after extinction, when the CS is given without the anticipated reward.

Question 7.9

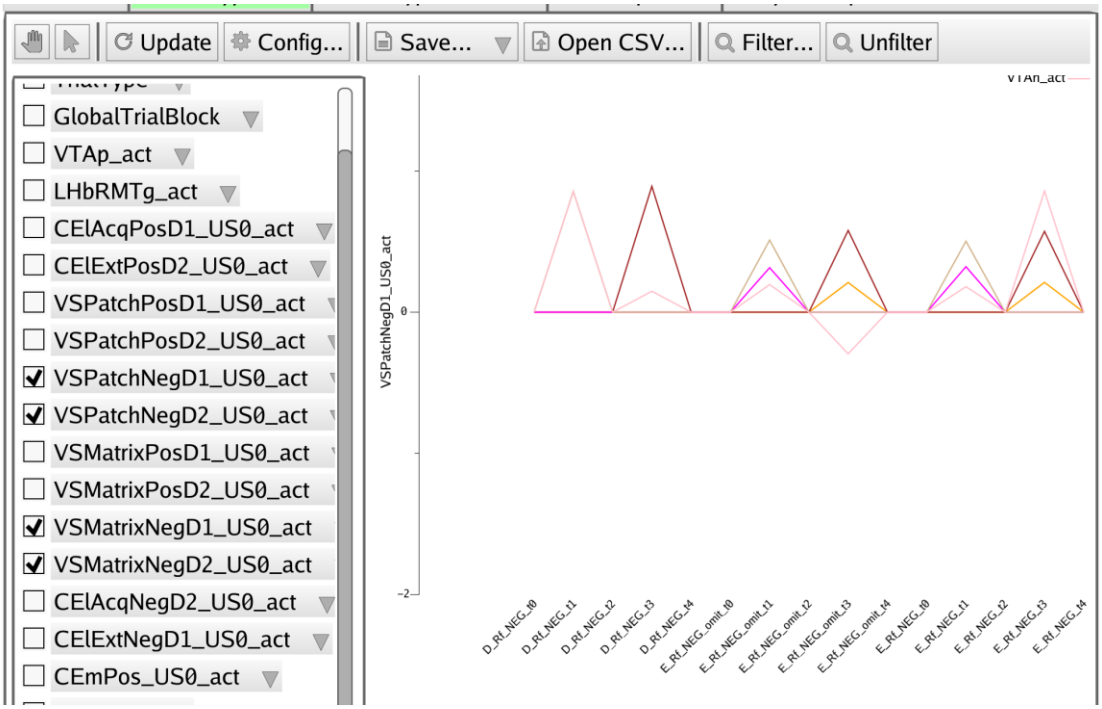


Evolutionarily speaking, a distinct extinction mechanism is better than the original learning's erasure-type process because it enables a more flexible and adaptive response to shifting environmental circumstances. Extinction learning's sensitivity to context

provides an important means by which organisms may distinguish between various circumstances and modify their behavior accordingly.

For an organism to survive, it must be able to establish links between stimuli and results since this allows the organism to anticipate and react to environmental occurrences. The predictive value of linkages may alter over time due to the dynamic nature of the environment. Extinction learning gives organisms a way to update these linkages while retaining the original knowledge, enabling them to adjust to changing conditions.

Question 7.9a



Separate learning pathways for aversive and appetitive primary outcomes would be better from an evolutionary point of view, as this would allow for specialized processing that handles the particular difficulties presented by both positive and negative environmental events. This can be adapted to ecological priorities and specialization for different learning needs.

If the positive responses to primary unpleasant outcomes in the VTAn layer were to be sent to downstream units that also receive signals from the VTAp, this could provide a difficulty for dopamine signaling. This is due to the fact that dopamine functions as an essential signal for behavioral adaptability and reinforcement learning. The correct assignment of valence to stimuli may be hampered if positive responses to unpleasant outcomes in VTAn were directly transmitted to downstream units.

Optional Question

The interleaving of A_Rf trials with AX trials allows the network to simultaneously learn both the positive association of stimulus A with reward and the negative

association of the conditioned inhibitor X with reward omission. And interleaving helps create a conflict between the positive and negative associations. Leading to the observed partial dopamine burst to the A stimulus when presented alone.

Dopamine neurons respond to the discrepancy between expected and actual rewards. In the case of interleaved A_Rf trials, the network expects a reward (positive association with A), but since reward is omitted during AX trials, there is a negative reward prediction error, leading to a dopamine burst.