A Neurocomputational Account of Catalepsy Sensitization Induced by D2-Receptor-Blockade in Rats: Context-Dependency, Extinction and Renewal:

Supplemental Material

Thomas V. Wiecki[1,2], Katrin Riedinger[3], Andreas von Ameln-Mayerhofer[3],

Werner J. Schmidt[3†], Michael J. Frank[4*]

[1] Wilhelm-Schickard-Institute for Computer Science, University of Tübingen

[2] Max Planck Institute for Biological Cybernetics, Tübingen

[3] Dept of Neuropharmacology, University of Tübingen

[4] Dept of Psychology and Program in Neuroscience, University of Arizona

[†] W.J. Schmidt deceased

[*]Corresponding author: mfrank@u.arizona.edu

January 31, 2009

Main changed parameters from original model

Layer	Param.	Orig. Val.	New Val.	Param	Orig. Val	New Val
Striatum	dt.vm	0.05	0.02	dt.net	0.04	0.7
	Inhibition	kwta	unit	$\overline{g_i}$	1	9
	$\overline{g_l}$	1	0.65	$\overline{g_i}$	1.2	0.1
Premotor Cort.	k_{hebb}	0.01	0.1			
Input→Striatum	lrate	0.001	0.003	k_{hebb}	0.01	0.4

Table 1: Summary of main changed parameters from the original model (Frank, 2006). Parameters not listed are the same as those in the original model and can be obtained from there. All changed parameters in the model striatum are simply due to the switch from the previous abstract kWTA inhibitory function to the the use of simulated inhibitory interneurons (see text), necessitating a change in inhibitory conductances. Further, the projections from premotor cortex to striatum were divided in two, one projects only to Go neurons, the other projects only to NoGo neurons. NoGo projections were set to be four times stronger than Go projections to roughly approximate empirical data

APPENDIX

The model can be obtained by emailing the first author at wiecki@tuebingen.mpg.de.

Implementation details

Like the original Frank (2006) model, this model is implemented using a reinforcement learning version of the Leabra framework (O'Reilly 1996; O'Reilly and Munakata 2000). Leabra uses point neurons with excitatory, inhibitory, and leak conductances contributing to an integrated membrane potential, which is then thresholded and transformed via an $\frac{x}{x+1}$ sigmoidal function to produce a rate code output communicated to other neurons (discrete spiking can also be used, but produces noisier results).

Because we are only interested in whether the model moves or does not (not in the selection between

competing responses, as in previous simulations), we simplify the original model by training and analyzing just a single response.

Parameter changes to the original model can be seen in Table 1. In addition, the current model uses a layer of simulated inhibitory interneurons (representing fast-spiking GABAergic interneurons) to regulate activation in the striatum, rather than the more abstract k-Winners-Take-All (kWTA) algorithm, which is described below. The reason for this change is purely technical (although it is also more biologically plausible). Specifically, because the abstract kWTA algorithm automatically regulates unit inhibitory conductances as a function of total excitatory input to the striatum, it does not support the ability to manipulate additional inhibitory effects of other input projections to the striatum. Thus in order to manipulate different levels of SNc $\rightarrow$ NoGo inhibitory projection strengths, to simulate differential D2 receptor occupation by haloperidol while leaving D1 receptor function intact, we eliminated the kWTA implementation and instead use a pool of 16 inhibitory interneurons to regulate striatal inhibitory conductances. (Parameters for inhibitory units are identical to those in inhibitory interneuron simulations of O'Reilly& Munakata (2000) Chapter 3.) As a result, D2 inhibitory effects can be additionally supported (interneurons and D2 projections contribute independently to inhibitory conductances). We note that this implementation with inhibitory interneurons does not critically alter any of our prior simulations of Go and NoGo learning as a function of overall DA manipulation, but is simply more computationally expensive.

The membrane potential V_m is a function of ionic conductances g with reversal (driving) potentials E as follows:

$$\triangle V_m(t) = \tau \sum_c g_c(t)\overline{g_c}(E_c - V_m(t)) \quad \text{(A-1)}$$

with 3 channels (c) corresponding to: e excitatory input; l leak current; and i inhibitory input. Following electrophysiological convention, the overall conductance is decomposed into a time-varying component

$g_c(t)$ computed as a function of the dynamic state of the model, and a constant $\overline{g_c}$ that controls the relative influence of the different conductances. The equilibrium potential can be written in a simplified form by setting the excitatory driving potential (E_e) to 1 and the leak and inhibitory driving potentials (E_l and E_i) of 0:

$$V_m^\infty = \frac{g_e \overline{g_e}}{g_e \overline{g_e} + g_l \overline{g_l} + g_i \overline{g_i}} \tag{A-2}$$

which shows that the neuron is computing a balance between excitation and the opposing forces of leak and inhibition. This equilibrium form of the equation can be understood in terms of a Bayesian decision making framework (O'Reilly and Munakata 2000). To simulate the basic effects of D1 and D2 receptor stimulation, D1 receptors activate an excitatory g_e current in striatal Go units, but with the level of DA affecting the gain of the unit's activation to simulate effects of D1 receptors on signal-to-noise (Frank 2005). To simulate D2 receptor inhibitory effects, the level of DA acted via the D2 receptor projection to increase inhibitory g_i currents on striatal NoGo units.

The excitatory net input/conductance $g_e(t)$ or η_j is computed as the proportion of open excitatory channels as a function of sending activations times the weight values:

$$\eta_j = g_e(t) = \langle x_i w_{ij} \rangle = \frac{1}{n} \sum_i x_i w_{ij} \tag{A-3}$$

The inhibitory conductance can either be computed by the kWTA function described in the next section or by modeling inhibitory interneurons. Leak is a constant.

Activation communicated to other cells (y_j) is a thresholded (Θ) sigmoidal function of the membrane po-

tential with gain parameter γ:

$$y_j(t) = \frac{1}{\left(1 + \frac{1}{\gamma[V_m(t) - \Theta]_+}\right)} \tag{A-4}$$

where $[x]_+$ is a threshold function that returns 0 if $x < 0$ and x if $x > 0$. To avoid dividing by 0 we assume $y_j(t) = 0$ if it returns 0. This activation is subject to scaling factors (wt_scale.abs and wt_scale.rel) which modify how much impact the projections have on the post-synaptic neurons.

Inhibition within and between layers

Inhibition between layers (i.e. for GABAergic projections between BG layers and striatal inhibitory interneurons) is achieved via simple unit inhibition, where the inhibitory current g_i for the unit is determined from the net input of the sending unit. For *within* layer lateral inhibition (used here in premotor cortex) , Leabra uses a kWTA (k-Winners-Take-All) function to achieve inhibitory competition among neurons within each layer (area). The kWTA function computes a uniform level of inhibitory current for all neurons in the layer, such that the k + 1th most excited unit within a layer is generally below its firing threshold, while the kth is typically above threshold. Activation dynamics similar to those produced by the kWTA function have been shown to result from simulated inhibitory interneurons that project both feedforward and feedback inhibition (O'Reilly and Munakata 2000). Thus, although the kWTA function is somewhat biologically implausible in its implementation (e.g., requiring global information about activation states and using sorting mechanisms), it provides a computationally effective approximation to biologically plausible inhibitory dynamics. kWTA is computed via a uniform level of inhibitory current for all neurons in the layer as follows:

$$g_i = g^{\Theta}_{k+1} + q(g^{\Theta}_k - g^{\Theta}_{k+1}) \tag{A-5}$$

where $0 < q < 1$ (0.25 default) is a parameter Θ for setting the inhibition between the upper bound of g_k and Θ . These boundary inhibition values are the lower bound of g_{k+1} computed as a function of the level of inhibition necessary to keep a unit right at threshold:

$$g_i = g^{\Theta}_{k+1} + q(g^{\Theta}_k - g^{\Theta}_{k+1}) \tag{A-6}$$

In the basic version of the kWTA function, which is relatively rigid about the kWTA constraint and is therefore used for output layers, g^{Θ}_k and g^{Θ}_{k+1} are set to the threshold inhibition value for the kth and k+1th most excited neurons, respectively. Thus, the inhibition is placed exactly to allow k neurons to be above threshold, and the remainder below threshold. For this version, the q parameter is almost always .25, allowing the kth unit to be sufficiently above the inhibitory threshold.

The premotor cortex uses the average-based kWTA version, g^{Θ}_k is the average g^{Θ}_i value for the top k most excited neurons, and g^{Θ}_{k+1} is the average of g^{Θ}_i for the remaining $n - k$ neurons. This version allows for more flexibility in the actual number of neurons active depending on the nature of the activation distribution in the layer and the value of the q parameter (which is typically .6), and is therefore used for hidden layers.

Catalepsy Measurement

Response times in the models are defined as the number of cycles (where one cycle is a single update of membrane potentials and firing rates of all units in the network as a function of current synaptic inputs) until the average activity in the two thalamic units exceeds a threshold of 0.475. Given that only one thalamic unit is active at a time (and in this case, there is only one valid response), this threshold corresponds to a single unit being 95% active. At this point, the bottom-up activity from thalamus to premotor cortex is sufficient to

preferentially facilitate one response and suppress the other in cortex. Response times were measured over 450 cycles. If the threshold was not reached during this time the response time was set to 450 cycles.

To quantify the measure of expected catalepsy in the model striatum, we use the following formula:

$$Catalepsy = \sum_{j=1}^{n_{NoGo}} act_j - \sum_{k=1}^{n_{Go}} act_k \quad \text{(A-7)}$$

where n_{NoGo} and n_{Go} represent the number of NoGo and Go neurons, respectively, and act_j and act_k the rate-coded activity of the jth NoGo and kth Go neuron, respectively. During NoGo-Go measurements the networks were only tested for 100 trials.

Each datapoint in the response time and the NoGo-Go plots is the average value of 30 networks, each initialized with a different set of random synaptic connection weights..

Learning

Synaptic connection weights were trained using a reinforcement learning version of Leabra (Frank 2005). The learning algorithm involves two phases, and is more biologically plausible than standard error back-propagation. In the minus phase, the model settles into activity states based on input stimuli and its synaptic weights, ultimately choosing a response. In the plus phase, the model resettles in the same manner, with the only difference being a change in simulated dopamine (the so-called 'phasic' component). Learning occurs during this second plus phase, as described below.

The original Frank (2005 2006) model simulated dopamine responses as an increase of SNc unit firing

from 0.5 to 1.0 for correct responses, and a decrease to zero in SNc firing for incorrect responses. Here we simulate rewards as an increase in phasic DA from 0.5 to 0.8 in the plus phase. The reason for this somewhat arbitrary change is that we expect the reward of an offset of an aversive stimulus to be smaller than that of e.g., a food reward. This specific value however is not crucial for the quality of our results. Nevertheless, although the simulated task is quite different from the original models' task, there is evidence for increased dopamine activity following the offset of an aversive stimulus (Jackson and Moghaddam 2004).

Leabra uses a combination of error-driven and Hebbian learning. The error-driven component is the symmetric midpoint version of the GeneRec algorithm (O'Reilly 1996), which is functionally equivalent to the deterministic Boltzmann machine and contrastive Hebbian learning (CHL), computing a simple difference of a pre and postsynaptic activation product across the two phases. For Hebbian learning, Leabra uses essentially the same learning rule used in competitive learning or mixtures-of-Gaussians which can be seen as a variant of the Oja normalization (Oja 1982). The error-driven and Hebbian learning components are combined additively at each connection to produce a net weight change.

The equation for the Hebbian weight change is:

$$\Delta_{hebb} w_{ij} = x_i^+ y_j^+ - y_j^+ w_{ij} = y_j^+ (x_i^+ - w_{ij}) \tag{A-8}$$

and for error-driven learning using CHL:

$$\Delta_{err} w_{ij} = (x_i^+ y_j^+) - (x_i^- y_j^-) \tag{A-9}$$

which is subject to a soft-weight bounding to keep within the 0-1 range:

$$\Delta_{sberr} w_{ij} = [\Delta_{err}]_+ (1 - w_{ij}) + [\Delta_{err}]_- w_{ij} \quad \text{(A-10)}$$

The two terms are then combined additively with a normalized mixing constant k_{hebb}:

$$\Delta w_{ij} = \epsilon [k_{hebb}(\Delta_{hebb}) + (1 - k_{hebb})(\Delta_{sberr})] \quad \text{(A-11)}$$

We found that for our simulations to produce reasonable results, the learning of the striatum needed a comparably high degree of Hebbian learning (k_{hebb} = 0.3) (relative to standard *Leabra* simulations in which the majority of the weight change is error-driven). While the exact learning rule in the striatum is still a topic of strong debate, there are multiple studies supporting a Hebbian-like rule (Kreitzer and Malenka 2005; Fino et al. 2005; Mahon et al. 2003). The reason that this is necessary in our simulations is that during extinction, NoGo units should not learn if they are prevented from being active (this is what leads to the non-extinguishable component). Hebbian learning ensures this to be the case. In contrast, if there was a greater proportion of error-driven learning, NoGo units would be more active during response selection (the minus phase) due to prior sensitization, and less active in the plus phase (during the phasic DA burst associated with reward / offset of the aversive stimulus). The difference in activity between the minus and plus phases would drive NoGo weights to further decrease during extinction, and would thus not produce as robust of a latent non-extinguishable sensitized component. We take this result as an indication for tuning the parameters of the original BG model, in which the k_{hebb} parameter was set to be the default, and somewhat arbitrary, value.

Novelty Model

Implemented Novelty Hypothesis Model

To implement this idea we again trained the model in the haloperidol mode for 60 epochs in context A. We additionally included simulated DA bursts in response to novelty. Thus, because the animal is still new to the experiment in the first days of testing, a simulated DA burst (high SNc activity) is presented together with the novel input context; the magnitude of this burst exponentially decays over trials as the environment becomes familiar (see Appendix for the specific form of the exponential). In epoch 60, the change to the novel context B is simulated by another DA burst, which again decays exponentially with experience. Recall, that in the reward prediction error model we assumed context B to be represented internally by a different set of input context units. Here we explore whether a similar catalepsy reduction can be observed if we assume the internal context cue to be the same in context A and B (which both contain overlapping environmental stimuli). All other implementation details are exactly as described for the reward prediction error model.

Results Catalepsy Sensitization

As shown in (Fig. 5 of main text), this model also produces a similar rise in cataleptic activity (i.e. sensitization), due to the same D2 blockade / NoGo learning mechanism. When the context changes in epoch 60 (simulated only by a novelty-induced DA burst and no change in input unit activity) we see a strong decrease in cataleptic activity (again matching the results by Klein and Schmidt (2003)). The renewed novelty-induced DA bursts potentiated Go activity and inhibited NoGo activity, leading to reduced cataleptic expression.

In sum, this novelty model provides a different explanation of the context-dependency observed in the empirical studies; an explanation that does not rely on the assumption that the different external context is also internally represented as a different context cue. This explanation also predicts that animals would

continue to show catalepsy expression if tested in a *familiar* context B, even if it had not been paired with haloperidol before (see below for further discussion).

Results Sensitized component

As described earlier, in the original Amtage and Schmidt (2003) study, a sensitized component was shown by placing the animal back in context A following extinction, and administering haloperidol. It was not tested whether the same sensitized component would be revealed in a different context. We showed above that the reward prediction error model predicts that this sensitized component is context-dependent. Here we show that the novelty hypothesis makes the opposite prediction.

To accomplish this, we first trained the model for 60 epochs in haloperidol model on context A with an exponentially decaying DA burst in trial 0, to simulate novelty. Then from epochs 60 to 100 we simulated the extinction training by switching the haloperidol model to the intact mode. Finally, in epochs 100 to 105 we tested for the sensitized component, by switching back to haloperidol mode.

Surprisingly, even though the test for the sensitized component was carried out in a novel context (simulated by a DA burst) we nevertheless observed an increase of cataleptic activity (Fig. 6 of main text). The reason for this is that NoGo units that had learned to increase their activity during sensitization were now re-activated in response to simulated haloperidol. The novelty DA burst was not able to completely inhibit this NoGo activity (given that D2 receptors were 90% blocked by the simulated haloperidol). Consequently, increased catalepsy expression is observed in the novel context, well above that seen at the end of extinction.

Thus, the novelty hypothesis predicts that the sensitized, non-extinguishable component is context-*independent*. This prediction contrasts with that of our reward prediction error model which predicts the sensitized component to be context-*dependent*.

Implemented Novelty DA burst

In our novelty models a dopamine burst is simulated during the onset of a novel stimulus. We thus added the value of an exponentially decaying function to the dopamine as follows:

$$act_{SNc} = \begin{cases} 1.0 & novel \\ act_{SNc} + exp(-i + 0.3) & otherwise \end{cases} \quad \text{(A-12)}$$

where i is the number of presentations of the stimulus since it was novel. While there is no direct evidence that the decay of DA activity in the SNc after repeated presentation of a novel stimulus is exponential, the specific form of the function is not crucial to our results and thus could have been chosen differently.

*

References

Amtage J, Schmidt WJ (2003) Context-dependent catalepsy intensification is due to classical conditioning and sensitization. Behav Pharmacol 14(7):563–567

Fino E, Glowinski J, Venance L (2005) Bidirectional activity-dependent plasticity at corticostriatal synapses. J Neurosci 25(49):11,279–11,287

Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. Journal of Cognitive Neuroscience 17:51–72

Frank MJ (2006) Hold your horses: A dynamic computational role for the subthalamic nucleus in decision making. Neural Networks 19:1120–1136

Jackson ME, Moghaddam B (2004) Stimulus-specific plasticity of prefrontal cortex dopamine neurotransmission. J Neurochem 88(6):1327–1334

Klein A, Schmidt WJ (2003) Catalepsy intensifies context-dependently irrespective of whether it is induced by intermittent or chronic dopamine deficiency. Behav Pharmacol 14(1):49–53

Kreitzer AC, Malenka RC (2005) Dopamine modulation of state-dependent endocannabinoid release and long-term depression in the striatum. J Neurosci 25(45):10,537–10,545

Mahon S, Casassus G, Mulle C, Charpier S (2003) Spike-dependent intrinsic plasticity increases firing probability in rat striatal neurons in vivo. J Physiol 550(Pt 3):947–959

Oja E (1982) Simplified neuron model as a principal component analyzer. Journal of Mathematical Biology 15(3):267–273

O'Reilly R (1996) Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. URL http://citeseer.ist.psu.edu/35164.html

O'Reilly RC, Munakata Y (2000) Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain. The MIT Press, Cambridge, MA