A Computational Model of Complex Skill Learning in Varied-Priority Training

[1]Wai-Tat Fu (wfu@illinois.edu), [1]Panying Rong, [1]Hyunkyu Lee, [1]Arthur F. Kramer, & [2]Ann M. Graybiel
[1]Beckman Institute of Advanced Science and Technology, University of Illinois, Urbana, IL 61801 USA
[2]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract

We reported a computational model of complex skill learning that captures the differential effects of Fixed Priority (FP) and Varied Priority (VP) training on complex skill learning. The model is developed based on learning mechanisms associated with the modular circuits linking Basal Ganglia, the prefrontal association cortex, and the pre-motor cortex during skill learning. Two forms of learning occur simultaneously. In discrimination learning, goal-directed actions are selected through recognition of external stimuli through the connections between the frontal cortex and the striatum, and is mediated by dopaminergic signals through a reinforcement learning mechanism. With practice, skill learning shifts from discrimination learning to Hebbian learning, which directly associates stimuli to responses by strengthening the connection between the prefrontal and pre-motor cortex. The model shows that FP training, in which all task components are equally weighted during training, leads to less flexible discrimination learning than VP training. The model explains why VP training benefits lower performance participants more, and why learning was more strongly correlated with the size of the striatum in VP than FP training.

Keywords: skill acquisition, skill transfer, multi-tasking, varied priority training, computational modeling

Introduction

Skill learning in complex, multi-tasking environments has long been an important cognitive science research topic. Although practice generally improves performance regardless of training methods, researchers have found that skill acquisition (practice time) alone is not sufficient to explain differences in effectiveness of these methods. It has been found that, for example, even though a training method may require a longer acquisition time, it may lead to better flexibility of skills to novel situations (Karmer, Larish, Strayer, 1995). In fact, in many practical domains (e.g., pilot training) the goal of training is seldom focused solely on training acquisition, as the trainee is often expected to perform in novel situations that require the skills to be flexibly deployed (Gopher, Weil, & Siegel, 1989).

Training Methods for Complex Skills

Among the different training methods for tasks with complex components, varied-priority (VP) training (e.g., Kramer et al., 1995) manipulates the relative emphasis of selected subcomponents in the multi-tasking environment while leaving the whole task intact (Gopher et al., 1989). Gopher et al. showed that systematically varying levels of priorities on attentional control through instruction and feedback could lead to better learning and performance in multi-tasking tasks. They argued that VP training enabled participants to explore different strategies and thus develop a better match between the requirements of the tasks and the efficiency of their efforts. As a result, VP training makes people better able to strategically allocate attention to multiple components of the task to comply with the change in emphases during training. In contrast, in fixed-priority (FP) training, all components are equally weighted, which was found to lead to learning of less flexible skills. Although benefits of VP training on global performance have been demonstrated through a number of studies, there is still a lack of understanding on the specifics of how it promotes learning and transfer. To the best of our knowledge, our model was the first that is developed at the neural computation level that explains the observed effects of the FP and VP training methods.

Neural Basis of Complex Skill Learning

Research shows that skill learning emerges as a result of the experience-dependent plasticity in the basal-ganglia-cortical neural circuits (e.g., Graybiel, 2008). Two major forms of learning are observed in these circuits. *Discrimination learning* allows recognition of pattern of stimuli and selection of correct responses. This form of learning requires executive processing of information at the prefrontal cortex (PFC) that guides the selection of actions, and is found to be mediated by external feedback. In contrast, *Hebbian stimulus-response (S-R) learning* allows direct association between stimuli and responses. This form of learning requires little executive processing, but often requires extensive training. Theories of skill acquisition often assumes that learning shifts from slow goal-direction behavior that requires executive processing to fast execution of S-R behavioral rules that requires less executive processing (e.g., Schneider & Shriffin, 1977).

Discrimination learning

During discrimination learning, goal-directed actions that require attentional function at the prefrontal cortex (PFC) are selected based on behavioral rules acquired through the declarative system (i.e., by following instructions in an experiment to associate a stimulus to a response). This form of learning involves the connections between the prefrontal cortex to the diverse set of spiny neurons in the striatum for pattern recognition computations (Houk & Wise, 1995) and the existence of relatively "private" feedback loops of connectivity from diverse cortical regions that converge onto those striatal spiny cells, via the pallidum and thalamus, and lead back to the frontal cortex (e.g., Amos, 2000; Kelly & Strick, 2004). Unlike neurons that learn through a Hebbian-like mechanism, spiny neurons are found to receive specialized inputs that appear to contain training signals from dopamine (DA) neurons (Fu & Anderson, 2006; Schultz, Dayan, & Montague, 1997; Schultz et al., 1995).

Hebbian S-R learning

Hebbian learning between the frontal cortex and the premotor cortex allows fast selection of responses tied to an environmental stimulus. Unlike learning at the striatum, Hebbian learning is often independent of the outcome of the responses, i.e., association is strengthened whenever the response is selected when the stimulus is perceived. During initial learning, because the correct S-R rules have not yet been learned, none of the responses will be activated. Instead, goal-directed behavior will guide the selection of the right response. With practice, the correct S-R rules are strengthened, which allow correct responses to be activated when the stimuli that tied to them are perceived.

The Role of the Striatum in Skill Learning

Research shows that the striatum are activated while performing tasks that require cognitive flexibility such as task switching and transfer to untrained tasks (Ragozzino et al. 2002; Meiran et al. 2004; Dahlin et al. 2008). PET studies in humans have shown that dopamine release and binding are increased in both of these striatal regions when subjects play a video game, and that greater dopamine binding is associated with better performance (Koepp et al. 1998). Erickson et al., (2010) shows that the differential size of the striatal regions predicts learning on an unfamiliar video game. They used magnetic resonance imaging (MRI)-based brain volumetry to measure striatal volumes of subjects with little previous video game experience before they received training on the classic Space Fortress video game (Mane & Donchin, 1989). They also compared the predictive value of the brain measures for different phases of learning including the initial acquisition period when performance was lowest but performance gains were highest. They found that individual structural differences in the striatum were effective predictors of procedural learning and cognitive flexibility, and were sensitive indicators of ventral-to-dorsal differences in striatal recruitment during learning. These findings suggest that changes in the striatum are predictive of learning effects observed during the video game. As we will show later, our model shows that discrimination learning at the striatum induced by different training methods can explain differences of their effectiveness in skill transfer.

The Model

The general structure of the model is shown in Figure 1. The activations of the neurons at the PFC (P) represent the different stimulus patterns perceived by the corresponding sensory cortical units (I). Neurons at the PFC are fully connected to the neurons at the association striatum (S), and the connection strength is changed through a dopamine-moderated discrimination learning mechanism (discussed next). The activated neurons at the striatum then send inhibitory signals to the globus pallidus (G), which send inhibitory signals to the thalamus (T). Neurons at the premotor cortex (M) are connected to both the thalamus and the PFC.

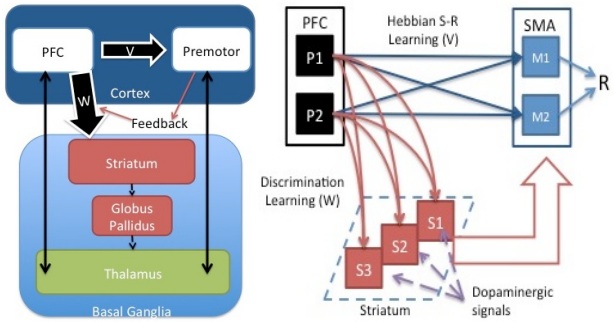

Figure 1. The basic structure of the model (left) and the two learning mechanisms (right) that allow the model to select the right actions. Detailed equations of the model can be found at the Appendix.

Goal-directed behavior is modeled by the connections from the PFC to the striatum, through the globus pallidus and thalamus, and eventually responses are selected in the premotor area. The thalamus is connected to both the PFC and the premotor area. The PFC is also directly connected to the premotor area, but the connections are weak initially, and thus selection of responses needs to go through the goal-directed path during the initial stage of learning.

Discrimination learning occurs at the connections between the PFC and the striatum (W), which is moderated by dopaminergic signals that reflect the valence (correctness) of the responses. Specifically, when the response was correct (a positive score) or incorrect (negative), the dopaminergic signal (D) would moderate learning of the connections (eqn 8-11). This form of dopamine-mediated learning is shown to resemble the reinforcement learning process that is extensively studied in machine learning (Fu & Anderson, 2006; Sutton & Barto, 1998). With practice and directed feedback, the model learns to select the correct responses when external stimuli are perceived (and interpreted by the PFC). The repeated firing of the correct S-R pairs strengthens the connections between the PFC and the premotor area through Hebbian learning. With enough practice, the connections become strong enough that the correct responses can be directly selected when the associated stimuli are perceived at the PFC, by-passing the slower subcortical path through the basal ganglia. The model therefore characterizes skill learning through the shift from discrimination learning through the basal ganglia to direct activation of the S-R rules at the cortex. The computational model was implemented by differential equations (see appendix) that simulate the activations of neurons in each brain structure shown in Figure 1.

Empirical Results

The goal of the model was to explain the functional characteristics of the model that explains the differences between VP and FP training. In this paper, we will focus on highlighting two major predictions of the model: (1) differences between VP and FP training for low and high

performance trainees, and (2) pre-existing size of striatum predicts later learning in VP more than in FP training. However, we will first describe empirical results demonstrating effects of VP and FP training schedules in an experiment that used a complex video game called *Space Fortress*.

The Space Fortress Game

The Space Fortress game was originally developed to study the acquisition of complex skills in fast-paced multi-tasking environments (Mane & Donchin, 1989). The main objective of the game was to maximize the total scores by shooting missiles at and destroying the space fortress, while maintaining a spaceship within a certain velocity limit and pre-specified boundaries on the screen (Figure 2). Missiles were fired from the spaceship. In addition to destroying the fortress, the participant had to protect his/her spaceship against damage from the fortress and mine. Participants used a joystick to control the spaceship, which flied in a frictionless environment. Participants not only needed to control the spaceship within boundaries, but also maintain its velocity within limits in the frictionless environment.

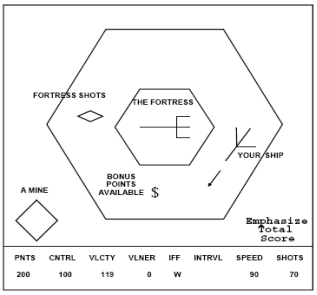

Figure 2. The Space Fortress game display

Participants were instructed to learn to control and maintain the spaceship within a particular range of velocity and a bounded area on the screen. These two subtasks were reflected by the velocity and control scores respectively, which were continuously updated on the screen. Participants also had to protect the spaceship from being hit by bombs emitted from the fortress and mines that periodically emerged on the screen. Participants could also shoot the mines to gain points. The four subscores: points (P), control (C), velocity (V), and speed (S) added up to the total scores, which were also continuously displayed on the screen.

In the Fixed Priority (FP) training condition, participants were instructed to give equal weight to the subscores throughout the sessions. In the Varied Priority (VP) training condition, participants were instructed to emphasize one of the four subscores in each game, and the emphasis changed throughout the sessions. Due to space limitation, we will focus on effects of the training conditions on the velocity subscore, which reflected how well the participants could successfully control the velocity of the spaceship. This subscore was also the most predictive of overall performance for all participants.

Effects of VP and FP training

Thirty-six participants from the University of Illinois community were randomly assigned to either the VP or FP training group. All participants completed the training in 10 consecutive days. Each day they did a 2-hour session, with each session consisting of 7 blocks. The first and last blocks are test blocks in which participants are required to emphasize total scores. There were 5 emphasis (practice) blocks between the test blocks. For the VP group, in each emphasis block participants were asked to emphasize some aspect of the game in the order of control, velocity, speed, points, and total score, and every other day, the reverse order. All emphasis conditions were communicated to participants by pop-up windows between sessions. Additionally, for the VP group, reminder text appeared at the corner of the display telling participants what they should be focusing on (see Figure 2). For the FP group, participants did the same amount of trials but are told to always emphasize total score.

To study whether pre-training performance difference might influence later learning, we performed a median split on the total scores of the first test block to identify the High (H) and Low (L) performance groups in each condition. Figure 3 shows the total scores for each group across the 20 test blocks. ANOVA showed a significant main effect of blocks ($F(19, 627) = 106.946$, $p<.001$), H-L ($F(19, 627) = 106.946$, $p<.001$), but not for conditions (FP vs. VP). There was a significant interaction between blocks and H-L ($F(19,627) = 3.891$, $p<.001$), and blocks and conditions ($F(19,627) = 1.745$, $p<.05$). Participants in the High and VP groups learned significantly faster than the Low and FP groups, respectively. The three-way interaction conditions x HL x blocks was marginally significant ($p=0.18$).

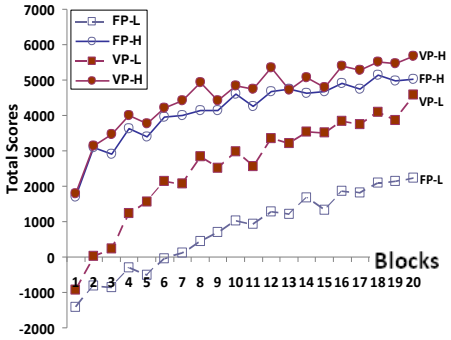

Figure 3. Average total scores across test blocks for the High (H) and Low (L) groups in each condition.

The results showed that, in general, VP training was more successful than FP training. Interestingly, the difference was larger in the Low performance group, in which participants started with a much lower score and was consistently lower throughout the 20 test blocks. In fact, Figure 3 shows that for the High performance group, participants in the VP condition were only slightly better than those in the FP conditions. However, for the Low performance group, the total scores for participants in the VP condition increased to almost the same level as the High performance group at the last block, but participants in the FP condition had a much lower total score even after 20 hours of training.

Size of the Striatum Predicts Learning

In the study by Erickson et al., (2010), they imaged the striatum with high-resolution MRI before the video game

training but after familiarization with the video game. They used an automated segmentation algorithm that employs a point distribution model from manual tracing of defined regions. After segmentation, the volume of each region was calculated based on voxel dimensions and adjusted for total ICV. The normalized volumes of the left and right caudate nucleus ranged from 3.80 to 7.43 cm^3 (mean = 5.33; SD = 0.85).

Analysis was conducted on performance across the entire 20-h training period, collapsed across both training groups, to determine whether striatal volumes were predictive of performance improvements in the Space Fortress game. Multiple regression analyses were conducted between change in performance and the volume of each region of the striatum, while including initial performance as covariates in the model. No correlation was found between striatum volumes and initial total scores in both groups, suggesting that initial performance of the task was not correlated with striatum volumes. However, striatum volumes significantly predicted change in performance when the groups were collapsed for the left ($F(2,33) = 4.00$; $p < 0.03$) and right ($F(2,33) = 3.94$; $p < 0.03$) caudate nucleus. The volumes of the dorsal striatal regions were positively correlated with training-induced performance improvements, and accounted for 23% of the variance in learning amounts across training. Analysis including the two training groups showed significant positive training group X volume interactions for the Points, Velocity, and Speed sub-scores. The dorsal striatal volumes were predictive of performance only for subjects in the VP group, but not for the FP group. In sum, results showed that the pre-existing volumes of the striatum predicted learning improvements *only* in the VP group, but not in the FP group. In other words, individuals with a larger striatum learned more effectively in VP training, but this benefit was not found in FP training. No such correlation was found between performance and other brain regions, such as the putamen and hippocampus.

The Simulations

For the present purpose, it suffices to use an abstract representation of the SF game and neuron activations in each brain region. Specifically, there were 1000 possible stimuli and 1000 possible responses, each represented by a vector with length 100. Each of the four subscores (velocity, speed, control, points) was considered a task component. Each response was randomly assigned a score in each task component (ranged from -4 to 4), such that the maximum total point for a response was 16 (4x4) and the minimum is -16 (-4x4). Each stimulus vector was directly fed to 100 neurons (P) in the PFC, which were fully connected to 100 neurons at the premotor cortex (M). The size of the striatum (S) varied from 20 to 100 neurons, each of which was connected to other structures (G & T) as shown in Figure 1. Based on the diffusion model of response decision (Ratcliff, 1978), when the integral of the difference between any two responses exceeded a threshold, the response with the largest activation would be selected (see eqn 6).

In addition to inputs from different regions, activations in each neuron were also decreased by two mechanisms (indicated by the negative terms, see eqn 1-5 at appendix): (a) lateral inhibition from neighboring neurons, and (b) decay of activation over time. Discrimination learning occurred at the connections (W) between PFC (P) and the striatum (S) (eqn 7). Discrimination learning depended on the strength of P, S, and the reward signal received (D). When the response was correct, the value of D would be positive (see eqn 10); when the response was incorrect, the value of D would be negative (see eqn 11). The connections that led to the correct response would then be reinforced based on the weighted sum of the task components (eqn 9). In VP training, the weights would change across blocks; in FP training the weights were all set to 0.25. This process was based on the reinforcement learning process that was shown to reflect the reward-based learning process at the basal ganglia (Fu & Anderson, 2006). Connections (V) between P and M were updated based on Hebbian learning mechanism (eqn 8), in which the strength of the connection is strengthened by an amount proportional to P x M.

Training and Testing of the Model

We randomly selected 500 stimuli for training in each session, and repeated the training for 20 sessions. In VP training, the weight for one task component was set to 0.85 and the rest set to 0.05 every 100 stimuli. We changed the parameter α to simulate the low (α = 0.01) and high (α = 0.05) performance groups, which controlled how fast it learns to select the correct responses.

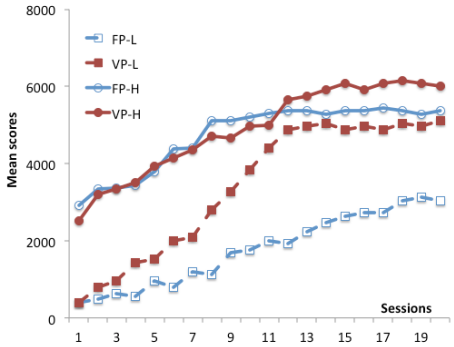

Figure 4. Mean scores of 100 simulations of the model across the 20 sessions.

Separate simulations were conducted for the low and high performance groups in VP and FP training. Figure 4 shows the mean proportion correct of the model during testing in each performance group in the VP and FP training conditions across 20 sessions. The simulation results show close resemblance to the empirical results shown in Figure 3. In particular, for the high performance groups, there was virtually no difference between VP and FP training. However, for the low performance groups, VP training led to much better learning than FP training.

In the model, the main reason why VP training was better than FP training in the low performance group was that discrimination learning was more effective in VP than FP training. Given that the learning rate (α) was lower, initial learning was equally slow for both VP and FP. However,

because the weight given to the emphasized task component was higher in VP training, feedback received was more effective because the value of D would be more *distinctive* when encoding the correctness of the response with respect to the emphasized task component (see eqn 9). In particular, when responses that were more important for a specific task component (e.g., a score of 4 in one component and 1 in the rest) and were emphasized during VP training, these responses would be reinforced more strongly in VP than FP training (another way to look at it is that in FP training, learning from feedback was diluted across components). In VP, feedback encoded by the dopaminergic signal is therefore more effective than in FP in guiding the model to strengthen the correct S-R pair through the feedback-driven reinforcement of the connections that activate the right set of neurons at the striatum, which eventually activate the correct responses at the premotor area. This subcortical pathway for response selection was then slowly transferred to the cortical-cortical S-R rules through the Hebbian learning mechanism (eqn 8).

In the high performance group, because the higher learning rate (α) compensated for the diluted feedback received in FP, the difference between VP and FP was reduced. As both groups achieved asymptotic performance, the difference between the two groups was not significant. However, further analysis did show that even in the high performance group, participants in VP training seemed more effective in learning sophisticated strategies than FP training, suggesting that VP training would more likely induce optimization of strategies with respect to each task component, while in FP training, responses that were generally good across components were learned. Due the space limitation, these analyses could not be included here.

To simulate effects of size of striatum on learning, the number of neurons in the striatum was increased from 20 to 100. Figure 5 shows the correlations between the percentage increase in performance of the model and the size of the striatum in the VP and FP training conditions. Consistent with empirical results, we found that the size of the striatum was positively correlated with performance improvement in VP training more than in FP training.

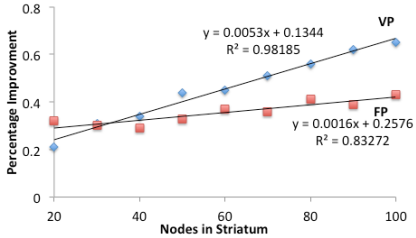

Figure 5. Correlations between percentage improvement (session 20 –1) and the number of nodes in the striatum.

In the model, the main reason why the size of the striatum predicted performance improvement more in VP training was because discrimination learning was more efficient with more neurons in the striatum. This effect was apparent from the model's perspective, considering the fact that the number of connections between the PFC and striatum would increase as the size of the striatum increased. A higher number of connections would naturally expand the capacity of the network to encode more S-R patterns.

The interesting question is why the model showed a stronger correlation in VP than FP training. This was again because *discrimination learning* was more effective in VP than in FP training. When different task components were emphasized, learning responses that led to higher score to that task component would be learned more efficiently in VP training. On the other hand, in FP training, learning of actions would more likely be based on their total scores, rather than their specific effects on each task component. The overall effect was that responses specific to certain task components were more likely encoded to different striatal neurons in the model in VP training (which would more likely lead to distinct responses selected), while FP training would more likely learn to select *generally good* responses. Because the granularity of the discrimination was higher in VP, learning would more likely be limited by the number of neurons at the striatum than in FP training (i.e., mapping between stimuli and responses was more sensitive to whether it was correct with respect to *each* task component in VP, thus practically creating another dimension in the mapping). This explained why performance improvement was more highly correlated with the size of the striatum in VP than FP training.

Conclusions and Discussion

We presented a computational model of complex skill learning at the level of neural computations between the prefrontal cortex, the basal ganglia, and the premotor cortex. The model successfully explained how VP and FP training induced different discrimination learning at the converging connections between the prefrontal cortex and the striatum, and how they eventually led to different effectiveness in overall learning. The model provided novel explanations to two major phenomena: (1) VP training benefits low performance participants more than FP training, and (2) the size of the striatum is highly correlated with performance improvement in VP but not in FP training.

In VP training, experiences of how different subcomponents were dynamically related to each other were learned more effectively than in FP training. Under FP training, participants received feedback based on the total score that represented the sum of subcomponents; while under VP training, participants received feedback that emphasized individual subcomponents. This difference led to more distributed and effective encoding of stimulus-response patterns at the striatum, which led to better overall training effectiveness.

The model demonstrated that the preexisting volumes of the striatum predicted performance improvement as subjects learning a complex video game, and the predictive power of the size of the striatum was much stronger in VP training. The model also captured these relations by showing that a larger striatum could accommodate more distributed S-R patterns experienced in VP training. In contrast, in FP training, discrimination learning would more likely select actions that were generally good across all task components,

and did not require as many neurons to encode the mapping. Thus, the size of the striatum was not a limiting factor in FP training. The model thus provided an explanation based on the interactions between training procedures and the computational characteristics of brain structures. The explanation was consistent with previous hypothesis that VP training could enhance coordination and integration of cognitive, motor, and perceptual operations, and allow more development of more flexible learning strategies. If VP training is more effective than FP for learning by capitalizing on the computational characteristics of basal ganglia-based circuits as a consequence, then this type of training could prove more useful for enhancing cognitive function in a number of applied settings.

References

Adams, J.A. (1987). Historical review and appraisal of research on learning, retention and transfer of human motor skills. *Psychological Bulletin 101*, 41-77.

Amos, A. (2000). A computational model of information processing in the frontal cortex and basal ganglia. *Journal of Cognitive Neuroscience, 12*, 505–519.

Dahlin E, Neely AS, Larsson A, Backman L, Nyberg L. 2008. Transfer of learning after updating training mediated by the striatum. *Science. 320,*1510--1512.

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to training of complex skills. *Acta Psychologica, 71*, 147-179

Graybiel, A. M. (2008). Habits, rituals and the evaluative brain. *Annual Review of Neuroscience, 31*, 359--387.

Erickson, Boot, Basak, Neider, Prakash, Voss, Graybiel, Simons, Fabiani, Gratton, Kramer (2010). Striatal volume predicts level of video game skill acquisition. *Cerebral Cortex, 293*, 1-9.

Fu, W.-T. & Anderson, J. R. (2006). From recurrent choice to skill learning: A reinforcement learning model. *Journal of Experimental Psychology: General, 135*, 184-206.

Houk, J. C., & Wise, S. P. (1995). Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: Their role in planning and controlling action. *Cerebral Cortex, 2*, 95–110.

Kelly, R. M., & Strick, P. L. (2004). Macro-architecture of basal ganglia loops with the cerebral cortex: Use of rabies virus to reveal multisynaptic circuits. Progress in Brain Research, 143, 449–459.

Koepp MJ, Gunn RN, Lawrence AD, Cunningham VJ, Dagher A, Jones T, Brooks DJ, Bench CJ, Grasby PM. 1998. Evidence for striatal dopamine release during a video game. *Nature. 393*, 266--268.

Kramer, A.F., Larish, J.F., & Strayer, D.L. (1995). Training for Attentional Control in Dual Task Settings: A Comparison of Young and Old Adults. *Journal of Experimental Psychology: Applied, 1* (10), 50-76.

Mane A., & Donchin, E. (1989). The space fortress game. *Acta Psychological, 71*, 17-22.

Meiran N, Friedman G, Yehene E. 2004. Parkinson's disease is associated with goal setting deficits during task switching. Brain Cogn. 54:260--262.

Ragozzino ME, Jih J, Tzavos A. 2002. Involvement of the dorsomedial striatum in behavioral flexibility: role of muscarinic cholinergic receptors. Brain Res. 953:205--214.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108.

Schneider, W. (1985). Training high performance skills: Fallacies and guidelines. *Human Factors 27*, 285-301.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127–190.

Schultz, W., Dayan, P., & Montague, P. R. (1997, March 14). A neural substrate of prediction and reward. Science, 275, 1593–1599.

Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. Cambridge, MA: MIT Press.

Appendix

The model was implemented as a set of differential equations shown below. The equation for one neuron at each brain structure was shown below. P=prefrontal cortex, S=Striatum, G=globus pallidus, T=thalams, M=premotor, D=dopaminergic signal, w=weights between prefrontal and striatum, v=weights between prefrontal and premotor, and i=input stimuli. α, β, and γ are free parameters that control the learning, and they were chosen to be equal (0.05) in each structure to minimize the number of parameters. The value of D depends on whether the response is correct (positive score) or incorrect (negative score).

$$\frac{dP_K}{dt} = (i_K + \alpha \cdot P_K)(1 - P_K) \cdot T_K - \beta(P_K) - \gamma \cdot P_J \quad (1)$$

$$\frac{dS_K}{dt} = [\sum_m w_{mK} P_J](1 - S_K) - \beta(S_K) - \gamma \cdot S_J \quad (2)$$

$$\frac{dG_K}{dt} = -\alpha(S_K G_K) - \beta(G_K) \quad (3)$$

$$\frac{dT_K}{dt} = -\alpha(S_K T_K) - \beta(T_K) + \alpha(T_K P_K) \quad (4)$$

$$\frac{dM_K}{dt} = \alpha(M_K) + [\sum_m v_{mK} P_K](1 - M_K) - \beta(M_K) - \gamma(M_J) \quad (5)$$

$$\Delta = \int (M_J - M_K) dt \quad (6)$$

$$w_{mK}(n+1) = w_{mK}(n) + \alpha \sum_t P_m \sum_t S_K \cdot D_K \cdot (1 - w_{mK}(n)) \quad (7)$$

$$v_{mK}(n+1) = v_{mK}(n) + \alpha \sum_t P_m \sum_t M_K \cdot (1 - v_{mK}(n)) \quad (8)$$

$$D_K = \sum_i weight(taskcomponent) \cdot D_K(taskcomponent) \quad (9)$$

When response is correct:

$$D_K(taskcomponent) = 1 - e^{\frac{S_K}{T}} / \sum_i e^{\frac{S_i}{T}} \quad (10)$$

When response is incorrect:

$$D_K(taskcomponent) = -e^{\frac{S_K}{T}} / \sum_i e^{\frac{S_i}{T}} \quad (11)$$