



Bridging across cognitive training and brain plasticity: a neurally inspired computational model of interactive skill learning

Wai-Tat Fu,^{1*} Hyunkyoo Lee,¹ Walter R. Boot² and Arthur F. Kramer¹

This article reviews recent empirical and brain imaging data on effects of cognitive training methods on complex interactive skill learning, and presents a neurally inspired computational model that characterizes the effects of these training methods. In particular, the article focuses on research that shows that variable priority training (VPT), which requires learners to shift their priorities to different task components during training, often leads to better acquisition and retention of skills than fixed priority training (FPT). However, there is only weak evidence that shows that VPT can enhance transfer of complex interactive skills to untrained situations. Brain imaging studies show that VPT leads to significantly lower activations and a higher reduction of activities in attentional control areas after training than FPT. Research also shows that the volume of the striatum predicts the learning effects, but only in VPT. The computational model, developed based on learning mechanisms at the neural level, bridges across the empirical and the brain imaging results by explaining the effects of VPT and FPT at both the behavioral and neural levels. The results were discussed in the context of previous findings on cognitive training. © 2012 John Wiley & Sons, Ltd.

How to cite this article:

WIREs Cogn Sci 2013, 4:225–236. doi: 10.1002/wcs.1214

INTRODUCTION

Skill learning in complex, interactive environments has long been an important cognitive science research topic.^{1–4} Decades of behavioral research has shown that although practice invariably improves performance, explicit training methods can more effectively enhance learning, retention, and transfer of skill.^{5–8} Recent research shows some evidence that cognitive training using simple laboratory tasks can improve specific basic cognitive operations and abilities, such as visual selective attention,⁹ working memory,¹⁰ or fluid intelligence.¹¹ On the other hand, research on cognitive training

using more complex interactive tasks (e.g., video games) shows mixed results—while training performance consistently increases, transfer of complex skills to untrained situations is limited.^{12,13} To further understand the effects of cognitive training, recent brain-imaging studies have focused on associating training outcomes to changes in brain activities.^{14,15} In particular, researchers are interested in understanding how different training methods may lead to changes in patterns of activities in different brain regions, and to identify biomarkers for effective acquisition, retention, and transfer of complex interactive skill.

Although there has been an increasingly large volume of empirical and brain imaging data, there is still a lack of an integrated theory that bridges across theories of cognitive training and brain plasticity to better predict the effectiveness of training strategies on interactive skill learning. The current focus article is

*Correspondence to: wfu@illinois.edu

¹Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL, USA

²Department of Psychology, Florida State University, Tallahassee, FL, USA

a first step toward such an integrated theory. First, we provide a brief review of recent advances in behavioral and brain imaging research on training of interactive skills that show how different training methods lead to changes in brain structure and function and behavioral performance. Second, we introduce a neurally inspired computational model that explains and predicts effectiveness of training methods based on mechanisms and representations at the neural level.

COGNITIVE TRAINING USING COMPUTER GAMES

The Space Fortress Game

One of the first attempts to systematically study effects of training methods on complex interactive skills was the learning strategies initiative.^{1,16} Under this initiative, a series of training strategies were outlined to enhance complex interactive skill learning using a multifaceted video game called space fortress (SF). Given that there has been a large volume of research using the same SF game, we briefly describe the general design of the game.

The SF game was designed such that learners need to master a large number of basic cognitive operations (as opposed to a single specific cognitive function, such as working memory, see e.g., Jaeggi et al.¹¹) by practicing when and how different cognitive operations should be combined and deployed. Learning to play the SF game, therefore, involves not only training on a specific cognitive function, but also the strategies for interleaving different operations, combining micro operations into macro operations, as well as the decision on when and how different operations should be used in dynamic situations. Training methods for this kind of complex interactive skills are, therefore, also more complex, and their effects may also differ from cognitive training methods that focus on a single cognitive function.

The main objective of the SF game is to destroy the SF, while maintaining a spaceship within a certain velocity limit and prespecified boundaries on the screen (Figure 1). These two subtasks are reflected by the velocity and control scores, respectively, which are continuously updated on the screen. Learners use a joystick to control the spaceship, which flies in a frictionless environment. Missiles are fired from the spaceship. In addition to destroying the fortress, the learner has to protect his/her spaceship against damage from the fortress and mines. Learners can also shoot the mines to gain points. In addition to tasks related to the fortress and mines, there is a constant monitoring task embedded in the game. Symbols

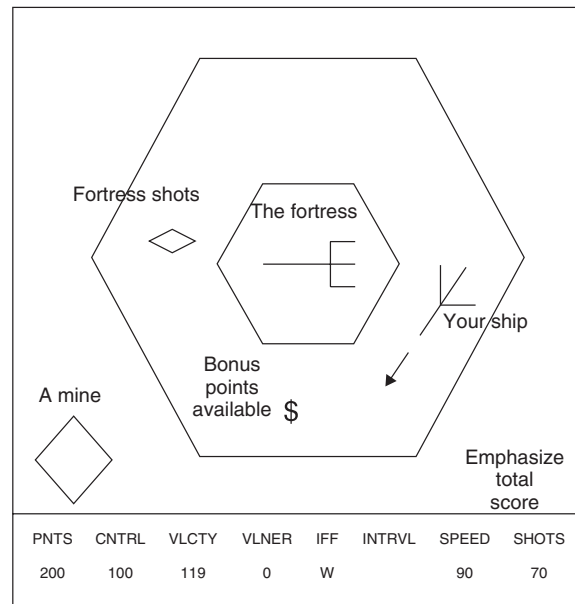


FIGURE 1 | The space fortress game display. See text for details of the game.

appear periodically below the fortress and whenever a dollar sign appears for the second time, players can use the mouse to either select bonus points or bonus missiles (which are a limited resource). However, if learners incorrectly identify the first dollar symbol as the second, they miss their opportunity to obtain a bonus when the second dollar sign does appear. Thus, learners are always encouraged to monitor this information.

The four subscores: points (*P*), control (*C*), velocity (*V*), and speed (*S*) add up to the total scores, which are also continuously displayed on the bottom of the screen (see Figure 1). Learners are asked to keep their ship within the two hexagons on the screen, and doing so increases the Control subscore. Flying the ship outside of the large hexagon or leaving the screen entirely (referred to as ‘going into hyperspace’) subtracts from the Control subscore. The velocity subscore rewards participants for going slowly and punishes participants for flying at high speeds. The speed subscore rewards/punishes participants for how quickly they deal with mines, and the points subscore rewards participants for shooting and destroying the fortress, but subtracts points for damage and destruction of the player’s ship. Typically, in fixed priority training (FPT), learners are instructed to give equal weight to the subscores (or just the total score) throughout the sessions. In variable priority training (VPT), learners are instructed to emphasize one of the four subscores in each game, and the emphasis changes throughout the sessions.

Given the complexity of the task, learning to effectively perform and integrate the various cognitive, perceptual, and motor operations requires at least 10–20 h of training. Because each of the components involves relatively different cognitive operations, analysis of changes in behavior across time not only can shed light on the effectiveness of training methods on improving performance in these components, but also the extent to which training can enhance basic cognitive operations and lead to transfer of these complex interactive skills to untrained situations.

A wide range of training strategies have been studied (see Ref 1). Part-task training involves principled decomposition of the complex task into functional components, which can be individually mastered by the learner. Whole-task training with VPT involves training on the integrated whole task with changing emphasis on the subcomponents of the task^{12,17–19}; while whole-task training with FPT involves equal emphasis of the subcomponents throughout the training. Finally, hybrid variable priority training (HVT) combines part-task training with VPT. Learners in HVT begin with a small number of components, but as they gradually master them, more components are added until the whole task is used (see Refs 17 and 20 for a comparison).

Behavioral Findings

Using the SF game, Gopher et al.¹⁷ showed that VPT could lead to better learning and performance in interactive tasks. They argued that VPT-enabled participants to explore different strategies and thus develop a better match between the requirements of the tasks and the efficiency of their efforts. Compared to FPT, VPT makes people better able to strategically allocate attention to multiple components of the task to comply with the change in emphases during training. Although benefits of VPT on global performance and some general transfer of skills to a new environment have been demonstrated through a number of studies¹⁷, there is still a lack of understanding on the specifics of whether and how it promotes transfer to untrained situations.^a

Recent research on cognitive training has led to a reexamination of various training strategies. VPT was found to be beneficial for dual-task performance,^{18,19} faster learning and higher level of mastery of complex skills,¹² and better working memory performance in older adults.²¹ Some studies found that VPT could lead to better transfer to untrained tasks¹⁸. A recent study by Lee et al.¹³ found that HVT (a variation of VPT) led to higher level of mastery of skills in the SF game than FPT, and the benefits of HVT remained even after 7 months. HVT was also found to be most

useful for participants who started training with lower performance, presumably because HVT allowed them to focus more on their weaknesses early on, such that their performance improved more in HVT than in FPT.

Contrary to previous claims¹⁷ that VPT could lead to transfer of skills to untrained tasks (often called ‘far transfer’, as opposed to ‘near transfer’, in which the trained and untrained tasks are similar or have functionally overlapping components), Lee et al.¹³ found that performance gained from VPT did not transfer to untrained tasks. The result was also in contrast to recent studies on cognitive training using simple laboratory tasks,¹¹ in which general increase in ability was observed after training for certain participants. On the other hand, the results were consistent with previous studies using complex interactive tasks,^{12,22,23} in which only near transfer was observed but not far transfer. The results not only suggest that far transfer is difficult to achieve, but also the fact that more research is needed to understand why certain training strategies using certain tasks lead to far transfer for some participants and not others. As we explain later, one possibility is to study the *cognitive representations* (a set of features that the person uses to encode the stimuli and responses in the environment) of the different task components as individuals are trained, as the same tasks could induce different cognitive representations and may influence skill learning and transfer.^{24,25} We elaborate on the role of cognitive representations when we introduce our computational model. First, we review recent neuroscience research on skill learning, which guides the construction of the computational model.

Brain Changes Induced by Training Strategies

In addition to behavioral research, recent brain imaging studies have shown how changes in brain activities induced by different training strategies differ. Using the SF task, Anderson et al.²⁶ compared brain changes during part-task and whole-task training. They found that individual learning gains could be predicted by critical regions of the brain involved when individual task components were learned. Most importantly, individual differences could be explained by the extent to which additional activation was observed when these regions already had a high level of activation. This finding was consistent with the cognitive load hypothesis,²⁷ which claims that learning effectiveness tends to decrease as brain activation reaches saturation (such that additional activation is limited). In other words, learners with

lower capacity may benefit more from part-task training as it requires less attentional demand, whereas learners with high capacity may benefit more from whole-task training as it provides more training opportunities per unit time.

Prakash et al.¹⁵ and Lee et al.¹⁴ compared brain activation changes induced by HVT and FPT as individuals learned to play the SF game over a 30-h period. Consistent with previous results, they found that individuals trained by HVT reached higher performance and showed less brain activation in areas related to visuospatial attention¹⁴ and top-down attentional control¹⁵ than those trained by FPT. In addition, HVT learners showed a significantly higher reduction of activation in these areas after training than FPT learners, along with reduced activation in the motor and sensory cortices and the posteromedial cortex. These findings suggest that VPT can lead to better mastery of skills by reducing the reliance on top-down attentional networks in the prefrontal cortices.

Given that the striatum plays a critical role in skill learning, Erickson et al.²⁸ imaged the striatum with high-resolution MRI before the SF game training using an automated segmentation algorithm, after which the volume of each region was calculated based on voxel dimensions. 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 SF game. Multiple regression analyses were conducted between changes 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 and right 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 \times volume interactions for the points, velocity, and speed subscores. The dorsal striatal volumes were predictive of performance only for subjects in the VPT group, but not for the FPT group. In sum, results showed that the preexisting volumes of the striatum predicted learning improvements *only* in the VPT group, but not in the FPT group. In other words, individuals with a larger

striatum learned more effectively in VPT, but this benefit was not found in FPT. No such correlation was found between performance and other brain regions, such as the putamen and hippocampus, supporting the critical role of the striatum in skill learning.

Summary of Results

As described above, results from behavioral and brain research provide support for the benefit of VPT on learning complex interactive skills. Brain imaging shows that VPT lead to higher reduction of activation in attentional control regions [the right middle frontal gyrus, right superior frontal gyrus, and the ventral medial prefrontal cortex (PFC)] compared to FPT,^{14,15} suggesting that VPT learners are able to perform better with less reliance on top-down attentional control. In addition, learners who have low initial performance tend to benefit more from part-task training and VPT,¹³ and individual learning gains were found to be limited by the cognitive load imposed by whole-task training.²⁶ To summarize, convergent evidence supports the following general characteristics of VPT. First, VPT is a more effective method for reaching the *automatic* stage of skill acquisition, in which stimulus-response associations (or *habits*) are selected and executed with less reliance on top-down attentional control. Second, one reason why VPT is more effective is because they tend to impose a lower cognitive load during training, such that individual components can be focused on as training emphasis switches.

A NEURALLY INSPIRED COMPUTATIONAL MODEL OF FLEXIBLE SKILL LEARNING

Researchers interested in developing integrated theories often rely on computational models to provide mechanistic accounts of behavior. To increase the biological plausibility of these models, researchers often rely on the large body of literature on computational properties of different brain structures to provide constraints on the models. These models are often referred to as *neurally inspired computational models*, as they attempt to provide multilevel explanations by simulating how computations in the brain can lead to specific behavioral patterns.^{29–31} We will adapt a similar approach to provide an integrated theory that bridges what we know about brain plasticity to effects of cognitive training by different training schedules.

Decades of research showed that skill learning emerges as a result of the experience-dependent

plasticity in the basal-ganglia-cortical neural circuits.^{29,32,33} Two major forms of learning are observed in these circuits. *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. *Discrimination learning* allows recognition of pattern of stimuli and selection of correct responses. This form of learning requires executive processing of information at the attentional control areas including the PFC that guides the selection of actions and is mediated by external feedback. Theories of skill acquisition often assumes that learning shifts from slow goal-directed behavior requiring executive processing to discriminate patterns of stimuli and the selection of correct responses to fast execution of S–R behavioral rules requiring less executive processing.^{34,35} The basal ganglia is often believed to be involved in the acquisition of these behavioral rules, in which activities at the caudate tend to increase during initial training and decline as the rules become ‘habitual’ (see Ref 36 for a review). We discuss how the co-occurrences and shift between these two learning processes may explain the benefits of VPT at both the neural and behavioral levels. However, we first briefly describe the neural components of these two processes, which inform the major components of the model.

Hebbian (Habit) Learning

Hebbian learning between the frontal cortex and the premotor cortex allows fast selection of responses tied to an environmental stimulus. Unlike learning in the striatum, Hebbian learning is often independent of the outcome of the responses, that is, 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.

Discrimination Learning

During discrimination learning, goal-directed actions that require attentional function at various brain regions including the 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 PFC to

the diverse set of spiny neurons in the striatum for pattern recognition computations³⁷ 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.^{38,39} 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.^{40,41}

The Role of the Striatum in Skill Learning

Research shows that the striatum is activated while performing tasks that require cognitive flexibility such as task switching and transfer to untrained tasks.^{33,42–44} PET studies in humans have shown that DA release and binding are increased in both these striatal regions when subjects play a video game, and that greater DA binding is associated with better performance.⁴⁵ As described above, Erickson et al.²⁸ shows that the differential size of the striatal regions predicts learning on an unfamiliar video game. 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 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 General Structure of the Model

The general structure of the model is shown in Figure 2. The activations of the neurons in the attentional control regions including the PFC (represented collectively in the figure as PFC) represent the different stimulus patterns perceived by the corresponding sensory cortical units (I). Neurons in the PFC are fully connected to the neurons in the association striatum (S), and the connection strength is changed through a DA-moderated discrimination learning mechanism (discussed next). The activated neurons in the striatum then send inhibitory signals to the globus pallidus (G), which send inhibitory signals to the thalamus (T). Neurons in the premotor cortex (M) are connected to both the thalamus and the PFC.

Goal-directed behavior is modeled by the connections from the PFC to the striatum, through the globus pallidus and thalamus, and eventually

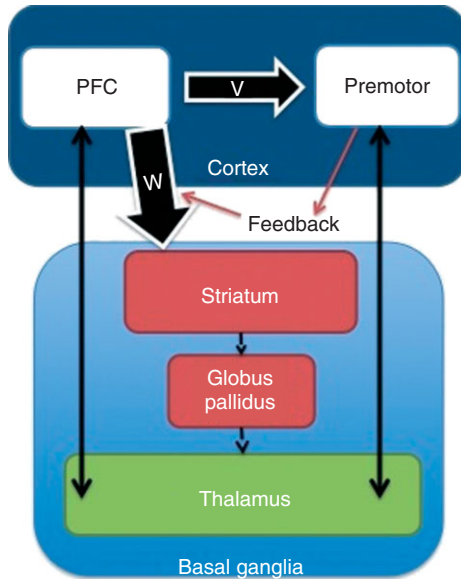


FIGURE 2 | The basic structure of the neutrally inspired computational model of flexible skill learning.

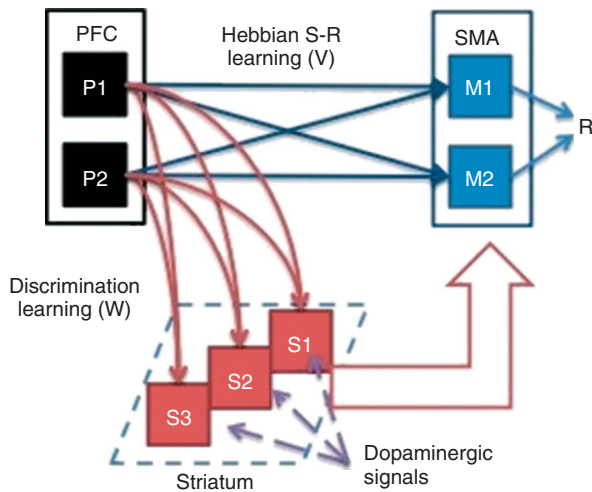


FIGURE 3 | The two learning mechanisms that allow the model to select the right actions. Detailed equations of the model can be found in the Appendix.

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 (Figure 3). Specifically, when the response was correct (a positive

score) or incorrect (negative), the dopaminergic signal (D) would moderate learning of the connections (Eqs (A8)–(A11)). This form of DA-mediated learning is shown to resemble the reinforcement learning process that is extensively studied in machine learning.^{40,46} 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 using a random walk model⁴⁷ 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 (a similar process on perceptual learning was modeled by Ashby et al.²⁹). The computational model was implemented by differential equations (see Appendix) that simulate the activations of neurons in each brain structure shown in Figure 2.

The Simulations

To illustrate computationally, the general effects of VPT and FPT on performance and brain activation changes, 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. The vectors represented by stimuli and responses were randomly generated by assigning either -1 or 1 to each element of the vectors (but no two stimuli or responses were identical), and the correct response vectors were randomly paired with the stimulus vectors. These numbers were chosen such that learning the right responses required a significant amount of learning, reflecting the amount of learning required by participants to master the task. On the other hand, we verified that increasing these numbers only prolonged the training periods of the model without impacting the patterns of results reported in this article. In other words, the general computational characteristics of the model are not sensitive to the specific sizes of the stimuli and responses.

In the simulation, each of the four subscores (velocity, speed, control, and 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 (4×4) and the minimum is -16 (-4×4). 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 and T) as shown in Figure 2. On the basis of the diffusion model of response decision,⁴⁷ when the integral of the difference between any two responses exceeded a threshold, the response with the largest activation would be selected (see Eq. (A6)).

In addition to inputs from different regions, activations in each neuron were also decreased by two mechanisms (indicated by the negative terms, see Eqs (A1)–(A5) in 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) (Eq. (A7)). 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 (Eq. (A10)); when the response was incorrect, the value of D would be negative (Eq. (A11)). The connections that led to the correct response would then be reinforced based on the weighted sum of the task components (Eq. (A9)). 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.⁴⁰ Connections (V) between P and M were updated based on Hebbian learning mechanism (Eq. (A8)), in which the strength of the connection is strengthened by an amount proportional to $P \times M$.

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 ($\alpha = 0.01$) and high ($\alpha = 0.05$) performance groups, which controlled how fast it learns to select the correct responses. These values were chosen to best fit the general patterns of the data we collected from the experiment. Note that there was no significant difference in the wide set of cognitive abilities that were tested between the two groups. Thus, the use of different learning rates to simulate the two groups was based on the finding that only the initial training performance was different between the two groups. We discuss alternative possibilities of the difference of the two groups in the discussion section.

Separate simulations were conducted for a low- and high-performance group in VPT and FPT, respectively. Figure 4 shows the mean proportion correct of the model during testing in each performance group in the VPT and FPT across 20 sessions. The simulation results show close resemblance to the empirical results by Boot et al.¹² shown on the left. In particular, for the high-performance groups, there was virtually no difference between VPT and FPT. However, for the low-performance groups, VPT led to much better learning than FPT.

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

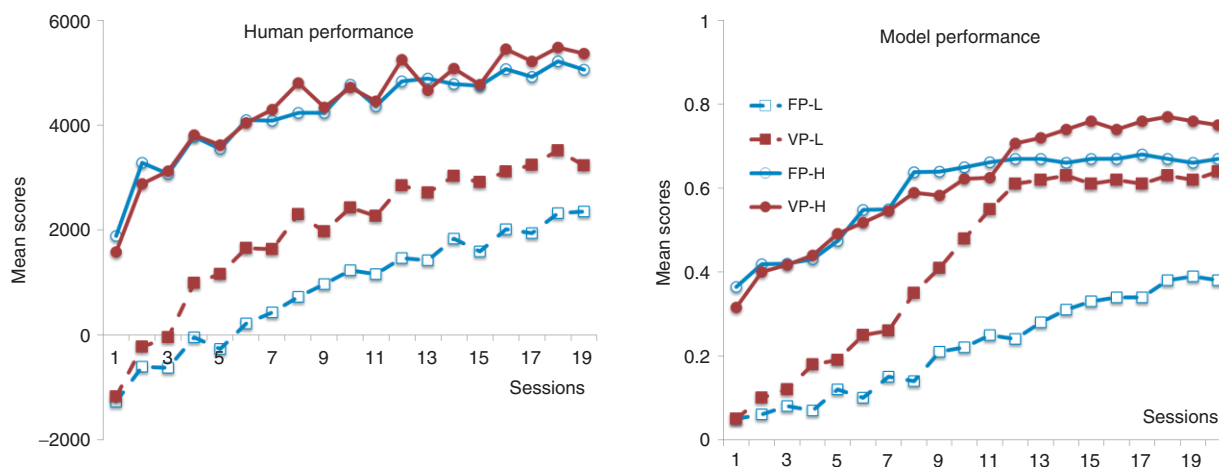


FIGURE 4 | Performance scores by human participants (left) and mean scores of 100 simulations of the model (right) across the 20 sessions. The model scores were normalized to 0 to 1. FP, fixed priority training; VP, varied priority training; L, low-performance group; H, high-performance group.

because the weight given to the emphasized task component was higher in VPT, 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 (Eq. (A9)). 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 in VPT, these responses would be reinforced more strongly in VPT than FPT (another way to look at it is that in FPT, learning from feedback was diluted across components). In VPT, feedback encoded by the dopaminergic signal is therefore more effective than in FPT to guide 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 more effectively transferred to the cortical-cortical S–R rules through the Hebbian learning mechanism (Eq. (A8)) in VPT than FPT. The results were consistent with the findings that brain activation in the attentional control regions (i.e., PFC in the model) reduced significantly more in VPT than FPT.

In the high-performance group, because the higher learning rate (α) compensated for the diluted feedback received in FPT, the difference between VPT and FPT 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 VPT seemed more effective in learning sophisticated strategies than FPT, suggesting that VPT would more likely induce optimization of strategies with respect to each task component, while in FPT, responses that were generally good across components were learned. Due to 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, the model shows that the size of the striatum is positively correlated with performance improvement in VPT more than in FPT.

In the model, the main reason why the size of the striatum predicted performance improvement more in VPT was because discrimination learning was more efficient with more neurons in the striatum. This effect was apparent from the model's perspective,

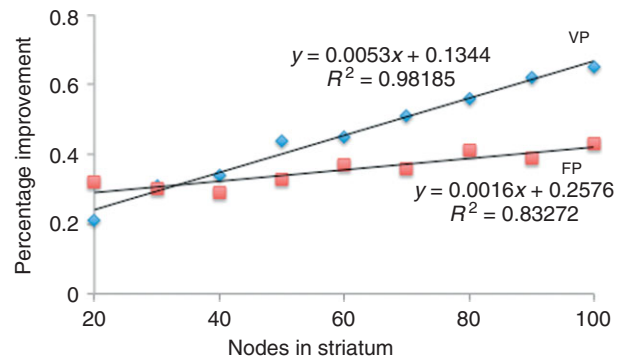


FIGURE 5 | Correlations between percentage improvement (sessions 20–100) and the number of nodes in the striatum.

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 VPT than FPT. This was again because *discrimination learning* was more effective in VPT than in FPT. When different task components were emphasized, learning responses that led to higher score to the task component would be learned more efficiently in VPT. On the other hand, in FPT, learning of actions would more likely be based on their total scores, rather than their specific effects on each task component (Eq. (A9)). The overall effect was that responses specific to certain task components were more likely encoded to different striatal neurons in the model in VPT (which would more likely lead to distinct responses selected), while FPT would more likely learn to select *generally good* responses. Because the granularity of the discrimination was higher in VPT, learning would more likely be limited by the number of neurons at the striatum than in FPT (i.e., mapping between stimuli and responses was more sensitive to whether it was correct with respect to *each* task component in VPT, thus practically creating another dimension in the mapping). This explained why performance improvement was more highly correlated with the size of the striatum in VPT than FPT.

CONCLUSIONS AND DISCUSSION

Consistent with previous results, VPT is found to be more effective than FPT for the acquisition of skills. Although there is some evidence that VPT enhances skill transfer to tasks that are functionally similar to some components of the trained tasks, broad transfer from trained to untrained tasks was

not generally observed. Learners with low level of initial performance before training tend to benefit more from VPT. Results are consistent with the saturation hypothesis in learning, in which the relatively high level of cognitive load imposed by FPT or whole-task training may create extra obstacles for learning of complex skills, while VPT may allow learners to focus on specific components to improve individual components of the skills.

To explain the effects of different training strategies, we developed a computational model of complex skill learning at the level of neural computations between the attentional control areas in the PFC, the basal ganglia, and the premotor cortex. The model successfully explained how VPT and FPT induced different discrimination learning at the converging connections between the PFC and the striatum, and how they eventually led to different effectiveness in overall learning. The model provided novel explanations to two major phenomena: (1) VPT benefits low-performance participants more than FPT, and (2) the size of the striatum is highly correlated with performance improvement in VPT but not in FPT.

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

In our model, we simulated the low and high-performance groups based on initial training performance (i.e., the learning rate). One may argue that differences in initial performance could be attributed to differences in preexisting ability in playing the game. However, we did not find significant differences in their general ability measures between the two groups, suggesting that the difference between the two groups lied at how well they learned during initial training. This is consistent with previous results¹¹ and prior meta-analysis,⁴⁸ which showed that performance gained during initial training predicted later training benefits. It is interesting to note, however, that, the current model predicts that initial training performance interacts with the training

method *and* preexisting capacity—i.e., a lower learning rate combined with VPT leads to better overall training, especially when the size of the striatum is large. Perhaps, because we are using a complex interactive task (as opposed to simple laboratory task such as the n-back memory task), learning multiple representations is more important for predicting training gain, and thus the size of the striatum is an important factor. On the other hand, a recent study has also found that reinforcement learning in the striatum can also explain why bilinguals have better executive function.⁴⁹ Thus, it is also possible that a larger striatum may also lead to broader improvement in general cognitive functions, although we did not observe this difference in the experiments using the SF game.

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 VPT. The model also captured these relations by showing that a larger striatum could accommodate more distributed cognitive representations of skills experienced in VPT. In contrast, in FPT, 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 FPT. 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 VPT could enhance coordination and integration of cognitive, motor, and perceptual operations, and allow more development of more flexible cognitive representations of the task environment. If VPT is more effective than FPT 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.

NOTE

^aAlthough Gopher et al. found that training in SF game transferred to new tasks, the similarities of the SF game and the new task were not well controlled in the study. See discussions in Ref¹³.

REFERENCES

1. Donchin E, Fabiani M, Sanders A. *The Learning Strategies Program: An Examination of the Strategies in Skill Acquisition*. Amsterdam: Elsevier; 1989.
2. Fu W-T, Anderson JR. Dual learning processes in interactive skill acquisition. *J Exp Psychol Appl* 2008, 14: 179–191.
3. Fu W-T, Chin J, Dong W, Liao V. Interactive skills and dual learning processes. In: Seel NM, ed. *Encyclopedia of the Sciences of Learning*. Springer-Verlag; 2011, 1612–1615.
4. Salvucci D, Taatgen NA. *The Multi-Tasking Mind*. New York, NY: Oxford University Press; 2010.
5. Carrier CA, Davidson G, Williams M. Selection of instructional options in a computer-based coordinate concept lesson. *Educ Commun Technol J* 1985, 33: 199–212.
6. Gopher D. Emphasis change as a training protocol for high-demand tasks. In: Kramer AF, Wiegmann DA, Kirlik A, eds. *Attention: From Theory to Practice*. New York, NY: Oxford University Press; 2007, 209–224.
7. Hannafin MJ. Guidelines for using locus of instructional control in the design of computer-assisted instruction. *J Instr Dev* 1984, 7:6–10.
8. Schmidt RA, Bjork RA. New conceptualizations of practice: common principles in three paradigms suggest new concepts for training. *Psychol Sci* 1992, 3:207–217.
9. Green CS, Bavelier D. Action video game modifies visual selective attention. *Nature* 2003, 423:534–537.
10. Klingberg T. Training and plasticity of working memory. *Trends Cogn Sci* 2010, 14:317–324.
11. Jaeggi SM, Buschkuhl M, Jonides J, Shah P. Short- and long-term benefits of cognitive training. *Proc Natl Acad Sci U S A* 2011, 108: 10081–10086.
12. Boot WR, Basak C, Erickson KI, Neider M, Simons DJ, Fabiani M, Gratton G, Voss MW, Prackash R, Lee H, et al. Transfer of skill engendered by complex task training under conditions of variable priority. *Acta Psychol* 2010, 135:349–357.
13. Lee H, Boot W, Basak C, Voss M, Prakash R, Neider M, Erickson K, Simons D, Fabiani M, Gratton G, et al. Performance gains from directed training do not transfer to untrained tasks. *Acta Psychol* 2012, 139:146–158.
14. Lee H, Voss M, Prakash R, Boot W, Vo L, Basak C, VanPatter M, Gratton G, Fabiani M, Kramer AF. Videogame training strategy-induced change in brain function during a complex visuomotor task. *Behav Brain Res* 2012, 232:348–357.
15. Prakash RS, De Leon AA, Mourany L, Lee H, Voss MW, Boot WR, Basak C, Fabiani M, Gratton G, Kramer AF. Examining neural correlates of skill acquisition in a complex videogame training program. *Front Hum Neurosci* 2012, 6:115.
16. Mane A, Donchin, E. The space fortress game. *Acta Psychol* 1989, 71:17–22.
17. Gopher D, Weil M, Siegel D. Practice under changing priorities: an approach to the training of complex skills. *Acta Psychol* 1989, 71: 147–177.
18. Kramer AF, Larish JF, Strayer DL. Training for attentional control in dual task settings: a comparison of young and old adults. *J Exp Psychol Appl* 1995, 1:50–76.
19. Kramer AF, Larish JL, Weber TA, Bardell L. Training for executive control: task coordination strategies and aging. In: Gopher D, Koriat A, eds. *Attention and Performance XVII: Cognitive Regulation of Performance: Interaction of Theory and Application*. Cambridge, MA: The MIT Press; 1999, 617–652.
20. Fabiani M, Buckley J, Gratton G, Coles MGH, Donchin E, Logie R. The training of complex task performance. *Acta Psychol* 1989, 71:259–299.
21. Stern Y, Blumen HM, Rich LW, Richards A, Herzberg G, Gopher D. Space Fortress game training and executive control in older adults: a pilot intervention. *Aging Neuropsychol Cogn* 2011, 18: 653–677.
22. Boot WR, Kramer AF, Simons DJ, Fabiani M, Gratton G. The effects of video game playing on attention, memory, and executive control. *Acta Psychol* 2008, 129:387–398.
23. Owen AM, Hampshire A, Grahn JA, Stenton R, Dajani S, Burns AS. Putting brain training to the test. *Nature* 2010, 465:775–778.
24. Kotovsky K, Hayes JR, Simon HA. Why are some problems hard? Evidence from the Tower of Hanoi. *Cogn Psychol* 1985, 17:248–294.
25. Rumelhart DE, Norman DA. Representation in memory. In: Atkinson RC, Hemstein RJ, Lindsey G, Luce RD, eds. *Steven's Handbook of Experimental Psychology*. 2nd ed. New York, NY: John Wileys & Sons; 1988, 511–587.
26. Anderson JR, Bothell D, Fincham J, Anderson AR, Poole B, Qin Y. Brain regions engaged by part- and whole-task performance in a video game: a model-based test of the decomposition hypothesis. *J Cogn Neurosci* 2011, 23:3983–3997.
27. Sweller J, van Merriënboer J, Paas F. Cognitive architecture and instructional design. *Educ Psychol Rev* 1998, 10:251–296.
28. Erickson KI, Boot WR, Basak C, Neider MB, Prakash RS, Voss MW, Graybiel AM, Simons DJ, Fabiani M, Gratton G, et al. Striatal volume predicts level of video game skill acquisition. *Cereb Cortex* 2010, 29:3:1–9.
29. Ashby FG, Ennis JM, Spiering BJ. A neurobiological theory of automaticity in perceptual categorization. *Psychol Rev* 2007, 114:632–656.
30. Wang H, Fan J. Human attentional networks: a connectionist model. *J Cogn Neurosci* 2007, 16: 1678–1689.

31. Wang H, Liu X, Fan J. Cognitive control in majority search: a computational modeling approach. *Front Hum Neurosci* 2011, 5:16.
32. Graybiel AM. Habits, rituals and the evaluative brain. *Ann Rev Neurosci* 2008, 31: 359–387.
33. Stocco A, Lebiere C, Anderson JR. Conditional routing of information to the cortex: a model of the basal ganglia's role in cognitive coordination. *Psychol Rev* 2010, 117: 540–557.
34. Fitts PM, Posner MI. *Human Performance*. Oxford: Brooks/Cole; 1967.
35. Schneider W, Shiffrin RM. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychol Rev* 1977, 84:127–190.
36. Seger CA, Spiering BJ. A critical review of habit learning and the basal ganglia. *Front Syst Neurosci* 2011, 5:1–9.
37. Houk JC, Wise SP. Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. *Cereb Cortex* 1995, 2:95–110.
38. Amos A. A computational model of information processing in the frontal cortex and basal ganglia. *J Cogn Neurosci* 2000, 12:505–519.
39. Kelly RM, Strick PL. Macro-architecture of basal ganglia loops with the cerebral cortex: use of rabies virus to reveal multisynaptic circuits. *Prog Brain Res* 2004, 143:449–459.
40. Fu W-T, Anderson JR. From recurrent choice to skill learning: a reinforcement learning model. *J Exp Psychol Gen* 2006, 135:184–206.
41. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science* 1997, 275:1593–1599.
42. Ragozzino ME, Jih J, Tzavos A. Involvement of the dorsomedial striatum in behavioral flexibility: role of muscarinic cholinergic receptors. *Brain Res* 2002, 953: 205–214.
43. Meiran N, Friedman G, Yehene E. Parkinson's disease is associated with goal setting deficits during task switching. *Brain Cogn* 2004, 54:260–262.
44. Dahlin E, Neely AS, Larsson A, Backman L, Nyberg L. Transfer of learning after updating training mediated by the striatum. *Science* 2008, 320: 1510–1512.
45. Koeppe MJ, Gunn RN, Lawrence AD, Cunningham VJ, Dagher A, Jones T, Brooks DJ, Bench CJ, Grasby PM. Evidence for striatal dopamine release during a video game. *Nature* 1998, 393:266–268.
46. Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press; 1998.
47. Ratcliff R. A theory of memory retrieval. *Psychol Rev* 1978, 85:59–108.
48. Colquitt JA, LePine JA, Noe RA. Toward an integrative theory of training motivation: a meta-analytic path analysis of 20 years of research. *J Appl Psychol* 2000, 44:117–186.
49. Stocco A, Yamasaki B, Natsenko R, Prat C. Bilingual brain training: a neurobiological framework of how bilingual experience improves executive function. *Int J Bilingual* 2012, 1–26.

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 , thalamus; M , premotor; D , dopaminergic signal; w , weights between prefrontal and striatum; v , weights between prefrontal and premotor; 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). See text for details.

In Eq. (A1), the time derivative of the activation of a neuron in the PFC (left side of the equation) increases with input stimuli from the sensory cortex (i), its own activation (P), and activation from the thalamus (T). The $(1 - P_K)$ term is to account for saturation, such that the activation cannot exceed 1. The negative terms represent inhibition, which include self-inhibition (P_K) and lateral inhibition (P_J).

In Eq. (A2), the time derivative of the activation of a striatum neuron increases as it receives excitatory inputs from the PFC (P), which are weighted by the massive number of convergent connections (w). Discriminatory learning occurs through the updating of these weights (Eq. (A7)). It has the similar saturation, self-inhibition, and lateral inhibition terms as in Eq. (A1).

In Eq. (A3), the time derivative of the activation of the globus pallidus neuron receives inhibitory inputs from the striatum. It also has the self-inhibition and lateral inhibition terms.

In Eq. (A4), the time derivative of the activation of the thalamus neuron receives excitatory inputs from the prefrontal cortex (P) as well as self-excitation. It also receives inhibitory inputs from the striatum. It also has the self-inhibition and lateral inhibition terms.

In Eq. (A5), the time derivative of the activation of a premotor neuron receives self-excitation (M), as well as the massive connections from the PFC (v). Hebbian learning occurs in these weights (Eq. (A8)). It has the same saturation, self-inhibition, and lateral inhibition terms.

In Eq. (A6), the integral represents the accumulation of the differences between the activations of different premotor neurons. As soon as this integral exceeds a threshold, the action represented by the premotor neuron with the highest activation will be selected.

In Eq. (A7), each weight is updated by the sum of activations in the prefrontal and striatal neurons. However, the update is also moderated by the DA signals, which depends on feedback in the environment (Eq. (A9)). It also has a saturation term ($1 - w$) that prevents its value from exceeding 1. Learning rate is controlled by the parameter α .

In Eq. (A8), each weight is updated by the sum of activations in the prefrontal and premotor neurons. It also has a saturation term ($1 - v$) that prevents its value from exceeding 1. Learning rate is controlled by the parameter α .

In Eqs (A9)–(A11), the value of the DA signal is assumed to varied depending on the feedback received from the task component. In fixed-priority training, the weight for each component will be equal, but in varied-priority training, one of them will be higher than the others (see text for details).

$$\frac{dP_K}{dt} = (i_K + \alpha \cdot P_K)(1 - P_K) \cdot T_K - \beta(P_K) - \gamma \cdot P_J \quad (\text{A1})$$

$$\frac{dS_K}{dt} = \left[\sum_m w_{mK} P_J \right] (1 - S_K) - \beta(S_K) - \gamma \cdot S_J \quad (\text{A2})$$

$$\frac{dG_K}{dt} = -\alpha(S_K G_K) - \beta(G_K) \quad (\text{A3})$$

$$\frac{dT_K}{dt} = -\alpha(S_K T_K) - \beta(T_K) + \alpha(T_K P_K) \quad (\text{A4})$$

$$\frac{dM_K}{dt} = -\alpha(M_K) + \left[\sum_m v_{mK} P_K \right] (1 - M_K) - \beta(M_K) - \gamma(M_J) \quad (\text{A5})$$

$$\Delta = \int (M_J - M_K) dt \quad (\text{A6})$$

$$w_{mK}(n+1) = w_{mK}(n) + \alpha \sum_t P_m \sum_t S_K \cdot D_K \cdot (1 - w_{mK}(n)) \quad (\text{A7})$$

$$v_{mK}(n+1) = v_{mK}(n) + \alpha \sum P_m \sum M_K \cdot (1 - v_{mK}(n)) \quad (\text{A8})$$

$$D_K = \sum_i \text{weight}(\text{task component}) \cdot D_K(\text{task component}) \quad (\text{A9})$$

When response is correct:

$$D_K(\text{task component}) = 1 - \frac{e^{\frac{S_K}{T}}}{\sum_i e^{\frac{S_i}{T}}} \quad (\text{A10})$$

$$D_K(\text{task component}) = -\frac{e^{\frac{S_K}{T}}}{\sum_i e^{\frac{S_i}{T}}} \quad (\text{A11})$$