

A Model of Practice Related Shifts in the Locus of Brain Activity During Verbal Response Selection Tasks

Vijaykumar Gullapalli
Mechanical & Aerospace Engg.
Princeton University
Princeton, NJ 08544
vijay@phoenix.princeton.edu

Jack J Gelfand
Psychology Department
Princeton University
Princeton, NJ 08544
jjg@phoenix.princeton.edu

Douglas L T Rohde
Computer Science Department
Princeton University
Princeton, NJ 08544
dlrohde@phoenix.princeton.edu

Abstract

Recent Positron Emission Tomography (PET) and other studies have produced detailed information about the areas of the brain involved in word association tasks, their functional roles in learning word associations, and the changes in activity in these areas during learning. We present a dynamic neuronal model that replicates observed human cognitive behavior in learning word associations while satisfying salient neuroanatomical and neuropsychological constraints. The model captures the observed dynamics of cortico-thalamo-basal ganglionic loops.

Introduction

Although “practice makes perfect” has been a long-standing dictum of skill learning, it is only recently that some light has been shed on the brain mechanisms involved in the learning of skills with practice. By combining information about the functional anatomy of the brain obtained through Positron Emission Tomography (PET) with other anatomical and physiological data, researchers have been able to piece together a more complete picture of the brain mechanisms involved in learning specific cognitive and motor tasks (e.g., (Raichle et al., 1994; Wise & Houk, 1994)). This, in turn, has facilitated the construction of realistic models of these brain mechanisms.

Our focus in this paper is on modeling the brain mechanisms involved in iterative verbal response selection tasks while satisfying known anatomical and functional constraints. We present a heterogeneous dynamic neuronal model that integrates diverse areas of the brain at a systems level. Our model replicates both the high-level cognitive behavior and the micro-level neuroanatomical characteristics of the diverse brain circuits involved.

Human cognitive behavior in iterated verbal response selection tasks is a simple, yet interesting, example of learning with practice. When subjects are asked to respond, for example, with appropriate verbs for a visually presented list of nouns, repeated presentation of the list initially elicits varying responses to each noun, but with practice, stereotypic responses develop. Practiced responses also are produced faster than those in the naive condition. In addition to these basic characteristics, verbal response selection also shows dependence on several cognitive variables, which, when manipulated, can give rise to priming, masking, interference, and other cognitive phenomena.

Related Research

In a recent set of studies, Raichle et al. (1994) and others used PET scans to examine the differences in the functional anatomy of the brain during naive and practiced performance of verbal response selection. PET scans over repeated presentations showed a shift in brain activity from the anterior cingulate, the left prefrontal and left posterior temporal cortices, and the right cerebellar hemisphere in the naive condition to the sylvian-insular cortex bilaterally and the left medial extrastriate cortex after practice. Introduction of a novel stimulus after practice reactivated the regions active in the naive condition. Raichle et al. (1994) concluded that two distinct brain circuits were employed in verbal response generation, one for controlled selection of responses and the other for the production of learned or automatic responses.

Complementary to the broad overview provided by the above-mentioned PET studies of the functionality of the brain areas involved in response selection, more focused studies have yielded fairly detailed information about the neuroanatomy and neurophysiology of the areas involved. For example, while cortical areas in general represent information regarding the internal state of the organism or the external state of the environment (e.g., (Mountcastle, 1978)), there is evidence to suggest that the cingulate cortex specifically represents information pertaining to the task that is currently being performed (Pardo et al., 1990; Vogt et al., 1992). Similarly, it is suggested that the sylvian-insular cortex serves as an associative store for learned responses to stimuli (Mitz et al., 1991; Raichle et al., 1994).

Finally, the volume edited by Houk et al. (Houk et al., 1995) (see also, Houk and Wise (Houk & Wise, 1993)) presents considerable evidence to indicate that the basal ganglia are involved in the selection and latching of relevant aspects of the current cortical state in a task-dependent fashion. This volume also describes the projections between areas of the cortex, the basal ganglia, and the thalamus in detail. In devising our model, which is based on the modular architecture described by Houk and Wise (Houk & Wise, 1993), we have adhered to the above-mentioned and other known anatomical and functional constraints.

Neural Modeling of Verbal Response Selection

Although verbal response selection is a high-level cognitive behavior, our model is implemented at a neuron level in order to enable us to incorporate our knowledge of the neurophysiology of the brain areas participating in the task. We have used lumped models of the neurons in our implementation, primarily because little is known about the impact of sub-neuronal dynamics on the high-level cognitive behavior. However, one of our goals is to determine the appropriate level of detail needed to model all the relevant aspects of verbal response selection in humans.

The most important feature of our model is that it incorporates the dynamical characteristics of the brain circuits involved. This allows us to replicate the temporal characteristics of the observed cognitive behavior, which is essential for any meaningful study of the phenomenon of learning. It will also enable us to compare the dynamic behavior of the model in simulations with the temporal characteristics of neuronal activations revealed by ERP data.

Description of the Model

The architecture used for our preliminary model is depicted in Figure 1. A primary feature of this architecture is the presence of two processing streams. The controlled stream consists of the cingulate cortex, the cortical modules representing the input words, the frontal cortex, the basal ganglia, and the thalamus. There is also an automatic stream, which consists of the sylvian-insular cortex. In the PET studies, these were the areas whose excitation differed significantly in the naive and practiced performance of the response selection task (e.g., (Raichle et al., 1994)). Both these streams receive common inputs from the sensory areas and send outputs to the motor areas.

The cortical modules representing the anterior cingulate and the sensory and language areas are organized as columns of neurons that correspond to cortical columns, each functioning as a relatively coherent information processing unit as discussed by Mountcastle (1978). Each cortical module forms a distributed representation of some internal state of the organism or external state of the environment. Cortico-cortical interconnections formed through Hebbian learning (Hebb, 1949) make it possible for the cortical modules to develop robust representations.

The representations used clearly have a significant impact on the functioning of a model. In the current model, we used a distributed representation over the sensory and language cortical modules to encode stimulus words, with each module denoting a "feature" or "category" of words (e.g., "colors" or "verbs"). The rationale for this representation is presented in the next section. These cortical representations are input to the basal ganglia, the sylvian-insular cortex, and the frontal cortex. Additionally, the cingulate module stores a representation of the task (e.g., "generate a color response"), and provides it to the basal ganglia and the sylvian-insular cortex.

We postulate that there would be a projection from the cingulate to the cerebellum as well and this will be implemented in future work as described in the discussion section.

Phasic activity in the sensory/language cortical modules is passed on to the frontal cortex through direct projections as shown in Figure 1. In addition, the frontal cortex also has highly specific reciprocal projections with the thalamus, resulting in local cortico-thalamic loops that, when active, may sustain activity in frontal cortex neurons (Houk & Wise, 1993). These loops could be activated through selective disinhibition by the basal ganglia (for review, see (Chevalier & Deniau, 1990)), which acts as a pattern-recognizer as suggested by Houk and Wise (Houk & Wise, 1993). We have modeled these interactions of the frontal cortex, the thalamus and the basal ganglia, which together form the controlled response pathway. Based on inputs from the cortical modules and the cingulate, the basal ganglia selectively disinhibit the frontal cortex-thalamus loops corresponding to the word features appropriate for the task, resulting in sustained activation of these features in the frontal cortex. Thus the output of the controlled circuit is an appropriate word represented by the tonically selected features in the frontal cortex.

In parallel with the controlled circuit, the sylvian-insular cortex module, which forms the automatic circuit, produces a response associated with the cortical inputs. Because of a lack of concrete anatomical evidence at this point, we simply modeled the insular learning as a linear associative network. Although this is not biologically faithful, it does have the property of incrementally learning the correct output response based on examples given by the performance of the controlled pathway and thus allows us to observe the overall dynamics of the model. Depending upon previous experience, the response given by the associative network simulating the sylvian-insular cortex may not be appropriate for the specified task. Therefore there is a need for deciding which response generation circuit to rely on. As mentioned in the discussion, we hypothesize that this role is played by the cerebellum, which would inhibit the inappropriate response.

In the current implementation, learning of synaptic weights takes place in the intra- and inter-cortical module connections, and in the sylvian-insular cortex. Although the other connections are currently hard-wired, we plan to incorporate learning in both the cerebellum and the basal ganglia in order to replicate the development of their pattern matching abilities. Considerable evidence regarding the presence of training signals and their operation in these areas is already available to guide this work.

Representation of words

Our use of a distributed representation over the sensory and language cortical modular array (Wise & Houk, 1994) to encode stimuli is inspired by the functional anatomy of the cortex (e.g., (Mountcastle, 1978; Asanuma, 1975)). The general organization of cortical circuits appears to be in the form of a distributed set

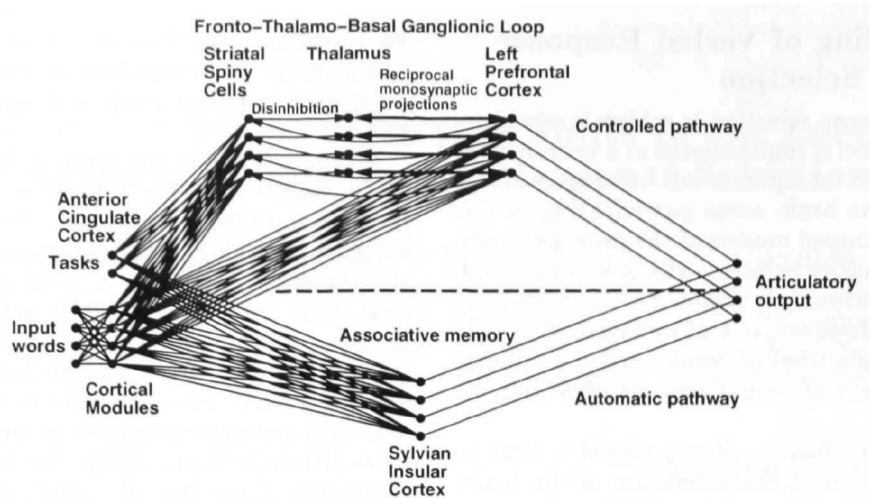


Figure 1: The architecture of the verbal response selection model.

of functionally specific regions or columns interactively involved in the execution of a given task. Each functionally specific region extracts from its inputs higher level information regarding a particular aspect of the task.

Sharing of processed information through reciprocal cortico-cortical projections between regions enables information extracted in one region to influence the processing of information in other regions concerned with the execution of the task. Cortical organization in columns with reciprocal projections between columns has been observed, for example, in the primary and secondary visual areas (Mountcastle, 1978), as well as in the motor cortex (Asanuma, 1975).

The cortical modular array (Wise & Houk, 1994) in our implementation corresponds to the local information processing regions of the cortex, with each module concerned with the representation of a “feature” or “category” of words. For example, a module might represent a color or colors associated with the stimulus word, or the fact that it is a verb. As a result, each word is represented as a distributed activation of the features associated with that word.

Currently, the nominal level of this distributed activation is predetermined for each stimulus word in our model. The levels of activation are given in Table 1. Thus, presentation of a stimulus word is effected by adding the activations indicated in the table to the corresponding neurons in the cortical modules. If no value is specified in the table for a neuron, its activity is not changed.

For example, when the stimulus “APPLE” is presented, (among others) the activations of the neurons representing “EAT” are increased by 0.5 and those of the neurons representing “SWEET” by 0.7, while the activations of the neurons representing “PET”, for instance, are not changed at all. These changes in activation values were selected to reflect the degree of association between the stimulus and the corresponding word. Thus, for instance, while “APPLE” is highly associated with both “RED” (1.0) and “FRUIT” (1.0), it is associated

to a lesser extent with “EAT” (0.5), and not at all with “BLACK” (0.0).

Two factors influence the temporal dynamics of these distributed activations. First, *shunting-type* lateral inhibition (e.g., (Pinter, 1983)) among the neurons in a cortical module results in a winner-take-all type selection of the feature represented by that module. Modeling neurons as leaky integrators with shunting-type inhibition leads to dynamic interactions between the new activity due to presentation of a stimulus and the previously existing activity of neurons in a module. These interactions play a significant role in determining what is represented in each module: previous strong activity of other features might inhibit weak new activity of a feature, resulting in suppression of a feature in the predetermined representation of the stimulus word. Alternatively, previous activation of the same feature might lead to priming of the new feature, increasing its prominence in the representation.

In contrast to this, the cortical modules attempt to maintain a coherent set of features in a representation by filling in missing features that were often active in the past in conjunction with those that are currently active. This is accomplished by mutually excitatory projections between modules in the cortical modular array whose strengths are determined via Hebbian learning (Hebb, 1949). Due to the dynamic nature of the representation of stimulus words, presenting the same word in different historical contexts can elicit different responses due to priming effects, much as in the case of human subjects.

Simulation Results

In this section, we present simulation results that demonstrate the features of our model. For these simulations, we selected a list of 20 words to represent in our model. Of these, six were used as stimuli, while the responses could be selected from all 20. The responses were classified into four groups, namely, stimuli, color names, verbs, and miscellaneous. A separate cortical module was used to represent words in each group. In Table 1, we present

Table 1: Representation of the stimulus words in the preliminary model as activations of cortical modular array neurons representing the features that comprise potential responses.

Stimulus	APPLE	BANANA	GRAPE	CAT	DOG	MOUSE
Potential Responses						
APPLE	1.0					
BANANA		1.0				
GRAPE			1.0			
CAT				1.0		0.2
DOG				0.5	1.0	
MOUSE				0.6		1.0
BLACK				0.8	0.7	0.9
BROWN			0.5	0.5	0.9	0.7
RED	1.0		0.3			
YELLOW		1.0				
BUY	0.9	0.9	0.9			
EAT	0.5	0.2				
FALL	0.8	0.8	0.7	0.2	0.3	0.1
RUN				0.8	0.9	0.9
HOUSE				0.3	0.3	0.2
FRUIT	1.0	0.8	0.9			
PET				1.0	1.0	1.0
SWEET	0.7	0.4	0.5			
SOUR	0.5	0.1	0.9			
TREE	0.4	0.1				

the (predetermined) activations of the words in each sensory/language cortical module when each of the stimulus words is presented.

The model captures the dynamics of cortico-thalamo-basal ganglionic loops suggested by Houk and Wise (1993) based on neurophysiological evidence gathered by several researchers (e.g., (Chevalier & Deniau, 1990; Fuster & Alexander, 1973; Goldman-Rakic & Friedman, 1991)). An example of this is presented in Figure 2. The first plot shows the activity of frontal cortex neurons when no task is specified. As can be seen, the initial activity of the neurons due to stimulus word presentation decays with time. When a “COLOR” task is specified, the activity of neurons representing a color associated with the stimulus word is sustained through selective activation of the corresponding loops, while the activity of the other neurons decays away. Similarly, when a “VERB” task is specified, activity of neurons representing actions associated with the stimulus word gets selectively sustained.

As reported by Raichle et al. (1994), the median response times of human subjects decreases significantly over successive blocks of presentation of the same set of stimuli. If a novel stimulus set is presented immediately following these repeated blocks, the response time returns to about the same level as in the naive condition for the original stimulus set. We ran a similar experiment with our system. We presented three stimuli, APPLE, DOG, and MOUSE, for 10 successive blocks, followed by 10 more blocks with the stimuli BANANA, GRAPE, and CAT. As illustrated in Figure 3, the response time of the system also decreased with repeated presentations of the same stimulus set for 10 blocks. Moreover, as with human subjects, presentation of a novel stimulus set caused

a significant increase in the response time, which further repetitions caused to decrease.

Discussion

The results given in the last section indicate that a model which is based upon the anatomical and functional features of the areas of the brain observed to participate in learning word association tasks does indeed exhibit the cognitive behaviour observed in humans. The model raises many important questions which suggest several directions for future research. Two major issues are discussed below.

A primary issue for investigation is the existence of dual (controlled and automatic) pathways for verbal response selection. Although the PET data from Raichle et al.’s study (1994) supports such a model, additional functional evidence obtained using alternative methods such as event related potentials (ERP) is needed to confirm or refute the conclusions of the PET study.

An important aspect requiring study is the functional role of the cerebellum. As mentioned above, increased right cerebellar activity has been observed in the naive condition of the word association task. We hypothesize that the cerebellum facilitates verbal response selection in two important ways: it arbitrates between the controlled and the automatic circuits, and it facilitates learning in the automatic pathway. This hypothesis is based on data from Feiz et al. (Feiz et al., 1992), who found that damage to the right cerebellar hemisphere of a human subject due to a stroke resulted the patient’s inability to select appropriate word association responses to words as well as his ability to learn at this task. We are in the process of incorporating this function of the cerebellum in our model.

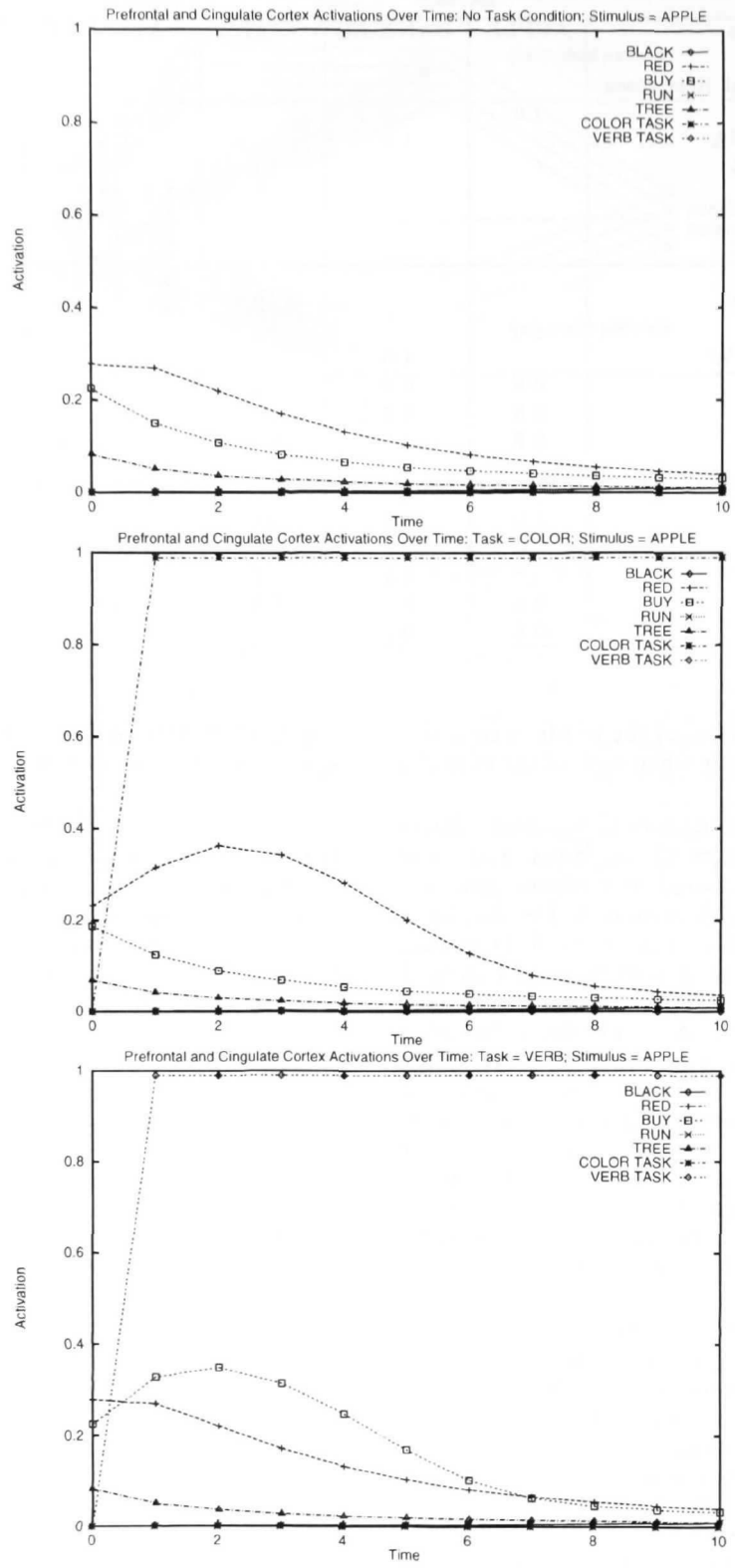


Figure 2: Example of selection of an appropriate response for different tasks by the controlled circuit. The color and verb task plots are activations of neurons in the cingulate cortex representing those tasks. The response activations (e.g., “RED”, “BUY”, “TREE”) are of neurons in the prefrontal cortex.

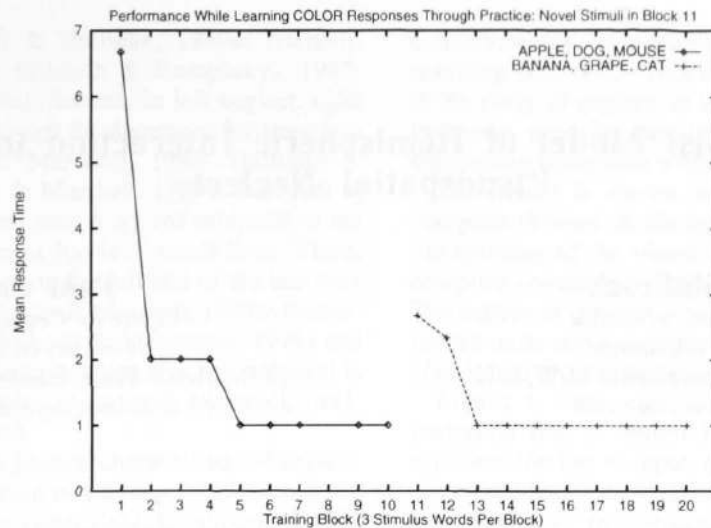


Figure 3: Learning verbal responses. Stimuli switched from (APPLE, DOG, MOUSE) to (BANANA, GRAPE, CAT) in block 11.

Acknowledgements

We would like to thank Prof. James C. Houk of Northwestern University for his guidance and many valuable suggestions. This work was partially supported by grants from the James S. McDonnell Foundation and the Office of Naval Research.

References

- Asanuma, H. (1975). Recent developments in the study of the columnar arrangements of neurons in the motor cortex. *Physiological Review*, *55*, 143–156.
- Chevalier, G. & Deniau, J. M. (1990). Disinhibition as a basic process in the expression of striatal functions. *Trends in Neuroscience*, *13*, 277–280.
- Feiz, J. A., Petersen, S. E., Cheney, M. K., & Raichle, M. E. (1992). Impaired non-motor learning and error detection associated with cerebellar damage. *Brain*, *115*, 155–178.
- Fuster, J. M. & Alexander, G. E. (1973). Firing changes in cells of the nucleus medialis dorsalis associated with the delayed response behavior. *Brain Research*, *61*, 79–81.
- Goldman-Rakic, P. S. & Friedman, H. R. (1991). The circuitry of working memory revealed by anatomy and metabolic imaging. In Levin, H. S., Eisenberg, H. M., & Benton, A. L. (Eds.), *Frontal Lobe Function and Dysfunction*, pages 72–91. New York: Oxford University Press.
- Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley.
- Houk, J. C., Davis, J. L., & Beiser, D. G. (1995). *Models of Information Processing in the Basal Ganglia*. Cambridge, Mass: MIT Press.
- Houk, J. C. & Wise, S. P. (1993). Outline for a theory of motor behavior: Involving cooperative actions of the cerebellum, basal ganglia, and cerebral cortex. In Rudomin, P., Arbib, M. A., Cervantes-Perez, F., & Romo, R. (Eds.), *Neuroscience: From Neural Networks to Artificial Intelligence. Research Notes in Neural Computation*, volume 4, pages 471–493. Berlin Heidelberg: Springer-Verlag.
- Mitz, A. R., Godschalk, M., & Wise, S. P. (1991). Learning-dependent neuronal activity in the premotor cortex: activity during the acquisition of conditional motor associations. *Journal of Neuroscience*, *11*, 1855–1872.
- Mountcastle, V. B. (1978). An organizing principle for cerebral function: the unit module and the distributed function. In Edelman, G. M. & Mountcastle, V. B. (Eds.), *The Mindful Brain*, pages 7–50. Cambridge, Mass: MIT Press.
- Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences, USA*, *87*, 256–259.
- Pinter, R. B. (1983). The electrophysiological bases for linear and for non-linear product term lateral inhibition and the consequences for wide field textured stimuli. *Journal of Theoretical Biology*, *105*, 233–243.
- Raichle, M. E., Feiz, J. A., Videen, T. O., MacLeod, A. M. K., Pardo, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during non-motor learning. *Cerebral Cortex*, *4*, 8–26.
- Vogt, B. A., Finch, D. M., & Olson, C. R. (1992). Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. *Cerebral Cortex*, *2*, 435–443.
- Wise, S. & Houk, J. C. (1994). Modular neuronal architecture for planning and controlling motor behavior. *Biological Communications of the Danish Royal Academy of Sciences*, *43*, 21–33.