

#### 4. General discussion: a unified model of cognitive control

Pfc has been widely discussed as being involved in tasks which require executive control, planning, WM, inhibition, decision making, and abstract thinking. All of the above tasks subserve a range of important functions in higher-level cognition (Gathercole, 1996; Logie, 1995; Logie & Gilhooly, 1998; Miyake & Shah, 1999; Richardson et al., 1996; D'Esposito et al., 1998; Owen, 1997; Smith & Jonides, 1999; O'Reilly et al., 2002). Prefronto-parietal and prefronto-temporal networks of brain regions are involved in the maintenance of task-set information (Gruber & von Cramon, 2001, 2003). Cognitive control is exerted via prefrontal top-down modulation of processing in domain specific sensory association areas in the parietal and temporal lobes. These descriptive accounts help to make sense of behavioral data, but nevertheless they provide little or no insight on how the cortical and subcortical areas involved in the task achieve these highly complex functions.

The present work suggests how such highly structured mechanism can be implemented in a biologically inspired neural system. This work should be contrasted to theoretical approaches, among which the so-called "central executive" and the SAS (Baddeley, 1986; Norman & Shallice, 1986; Shallice, 1982; Shallice, 1988), which do not have the necessary articulation needed for a modern theory of cognitive control. The model should also be contrasted and compared with the modeling studies discussed in this thesis, from which this work emerges not only as a biologically-motivated effort, but as an attempt to unify several aspects of cognition, learning and performance. The modeling studies reviewed in Chapter 2 (Braver and Cohen, 1999; Durstewitz and Seamans, 2002; Dreher and Burnod, 2002) show how achieving a good balance between biological plausibility and the range of behavior the model is trying to explain is a challenging task. The models described in this thesis have been designed with the aim to strike a balance between these constraints, achieving good results in terms of the amount of neurobiological and behavioral data explained while keeping the degree of complexity in the systems as low as possible.

The model has shown how complex behaviors can be learned by a "dummy" strategy based on trials and errors. This is a very important feature of the system, since it does not require an external teacher or a supervisor in order to shape the learning process. This is particularly important for an approach that aims at modeling biological autonomous systems, in which an external teacher is not always present and the system should most of the time rely on its own experience. Furthermore, external contingencies can vary as well as the timing, availability and magnitude of reinforcement. Therefore, a system that can dynamically and autonomously shape its behavior according to variable reinforcement contingencies has a clear ecological advantage.

One interesting feature of the model is the role of the Hippocampus. Both in the model and in experimental manipulations, the Hippocampus is crucial in the first stages of learning, whereas its

damage does not impair consolidation of memories when performed after a critical period (Kim et al., 1995). In the model, this effect is accounted for by the fact that the Hippocampus drives consolidation of LTM in neocortex, which then stores associations within cortico-cortical synapses. Nevertheless, some properties of the circuit will be impaired, namely the ability of the system to generate predictions on the exact timing of the reward (through the AT mechanism and NAc).

The model mirrors the main anatomical and physiological data which show that the hippocampus receives multimodal sensory information. The long study of place cells within the hippocampus proves that, at the very least, the hippocampus processes visuospatial information (Rotenberg and Muller, 1997). More recent data from the hippocampal electrophysiology literature indicate that other stimuli, like olfactory stimuli (Eichenbaum et al., 1987; Wood et al., 1999) and auditory stimuli (Edeline et al., 1988; Luntz-Leybman et al., 1992; Adams and Stevens, 1998), reach the hippocampus. The question which is still open is whether or not the hippocampus simply serves to relay this sensory information to another part of the fear circuit that serves a mnemonic and/or output function. However, experimental observations argue against a simple sensory role for the hippocampus: post-training lesions are not effective if delayed for a considerable time after training (Kim and Fanselow, 1992; Maren et al., 1997).

An interesting recent paper by Bailey et al. (2002) reports that manipulation of gamma-aminobutyric acid (GABA) **transmission within the hippocampus can cause a fear response prior to the administration of any footshock**. The GABA receptor agonist, RY024, caused both an increase in fear-related behavior **before** footshock administration and a **reduction in conditioned fear when animals were tested later off drug**. These results are an indirect confirmation of the importance of hippocampus in timing the behavioral response, a result which is clearly consistent with the model. In fact, the modeled hippocampus exerts a timed excitation on the NAc, which in turn inhibits VTA, thereby providing a timed signal for the control of the crucial DAergic pathway. A disruption of this adaptively-timed control might lead to the premature release of behavior, like in an anticipated fear response, as reported by Bailey et al. (2002).

A further interesting point is the analogy between the subdivision of negative/positive phases and the sleep/wake cycles typical of the majority of animal species. It would be interesting to explore the analogies between the biphasic structure of the model and the analogous biphasic functional mode expressed by biological nervous systems.

A final consideration concerns the “fragility” of the structures involved in cognitive control. As it is evident from the complexity of the model, balancing all the various components of the system is not an easy task. It is therefore not surprising that Pfc dysfunctions are involved in most pathologies of executive functions. Pfc damage has been associated with increased distractibility and perseveration (Damasio, 1985). A deficit in Pfc cortex functioning has been correlated with schizophrenia (Goldman-Rakic, 1995, Reid and Willshaw 1999). An underactive or impaired Pfc is believed to be responsible of

some of the major deficits seen in schizophrenia, namely thought process disturbances (Passingham, 1993; Goldman-Rakic, 1995). The linkage between an imbalanced DAergic system, a poor Pfc activation and some key symptoms of schizophrenia (Goldman-Rakic, 1995) is one of the major issues in schizophrenia research (Lidow et al., 1998; Braver and Cohen, 1999; Reid and Willshaw, 1999; Benes et al., 1999). Although the present model does not openly address any psychopathological data, several interesting results have emerged especially in the first set of simulations.

The simulations have shown, in fact, how DAergic hyper and hypo-activity cause different patterns of working memory (WM) impairment. DAergic hypoactivity prevents the instantiation of the WM pattern, causing the network to be driven by external input, and could be approximated to the increased distractibility shown by patients with a Pfc lesion, in particular the fact that hypofrontal patients are easily distractible by environmental events. On the other side, DAergic hyperactivity has a different family of associate deficits. An overactive DAergic system causes, in the model, the possibility of interference from “inappropriate” patterns into WM, or the expression of non-preponderant units activation due to the amplification of the recurrent connections in Pfc. Although these conclusions are surely interesting and promising, the results are still tentative and need further work.

The natural extensions of this thesis would be, therefore, to investigate how imbalances in the system can shape our understanding of psychopathology. These imbalances can represent the counterpart of ontogenetically, phylogenetically or stress-induced variation in a parameter space that characterize the model, whose architecture and functioning are greatly affected by the right choice of parameters.

Finally, it is important to stress how contemporary models of how the system *should be designed in order to perform its normal functioning* are lacking, whereas models of “abnormal” behavior are proliferating. Studying the pathology directly without attempting to understand how the brain gives rise to the normal behaviors that are eliminated or impaired during the disorder is, in my opinion, a hopeless endeavor. This work can be therefore considered as a small step towards the design of a biologically plausible model of higher cognitive functions.