Forthcoming in Spurrett, D., Kincaid, H., Ross, D. & Stephens, L. (eds.) (in preparation) Distributed Cognition and the Will: Individual Volition and Social Context, Cambridge, Mass.: MIT Press.

# The evolutionary origins of volition

Wayne Christensen
Philosophy, University of Kwazulu-Natal
wayne.christensen@ukzn.ac.za

## 1 A high order control basis for volition

It appears to be a straightforward implication of distributed cognition principles that there is no integrated executive control system (e.g. Brooks 1991, Clark 1997). If distributed cognition is taken as a credible paradigm for cognitive science this in turn presents a challenge to volition because the concept of volition assumes integrated information processing and action control. For instance the process of forming a goal should integrate information about the available action options. If the goal is acted upon these processes should control motor behavior. If there were no executive system then it would seem that processes of action selection and performance couldn't be functionally integrated in the right way. The apparently centralized decision and action control processes of volition would be an illusion arising from the competitive and cooperative interaction of many relatively simple cognitive systems. Here I will make a case that this conclusion is not well-founded. Prima facie it is not clear that distributed organization can achieve coherent functional activity when there are many complex interacting systems, there is high potential for interference between systems, and there is a need for focus. Resolving conflict and providing focus are key reasons why executive systems have been proposed (Baddeley 1986, Norman and Shallice 1986, Posner and Raichle 1994). This chapter develops an extended theoretical argument based on this idea, according to which selective pressures operating in the evolution of cognition favor high order control organization with a 'highest-order' control system that performs executive functions. The core elements of this architecture are presented in figure 1.

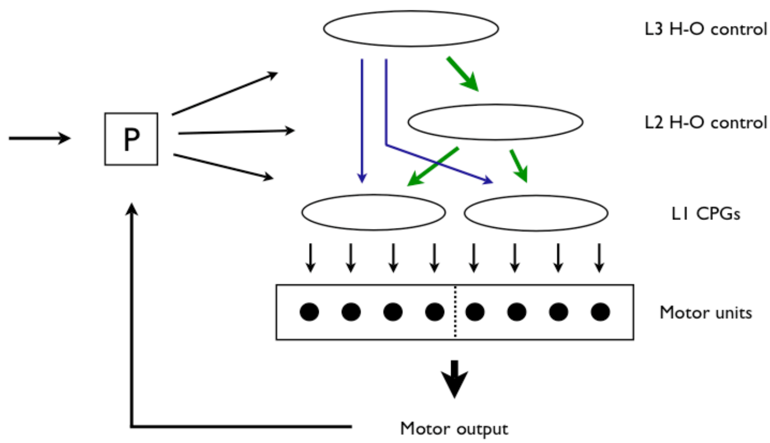


**Figure 1: High order control architecture.** Key properties: (i) Motor and perceptual systems have many degrees of freedom. (ii) The first level of control is provided by CPGs, which determine patterns of motor activation. (iii) High order control systems are differentially specialized for increasingly high order control problems in ascending order; lower order control systems provide constrained stereotypical control, higher order systems provide increasingly flexible high dimensional control. (iv) All control systems can access perceptual information directly and receive descending influence from higher systems. For simplicity only descending connections are shown, but ascending connections are assumed. (v) Perception-action loops are possible for each level of control. (vi) Higher level systems are only engaged as necessary. (vii) The highest-level system can exert top-down influence either via intermediate control systems or via direct control of level 1 controllers, permitting either coarse or fine-grained influence on motor control in varying circumstances. Three levels of control are shown, but in actual cases there will typically be more than this. CPGs: central pattern generators; H-O: high order. Cf. Swanson 2003a, figure 6.7 and Fuster 2004, figures 1 and 2.

According to the high order control model control competency is distributed across multiple systems but systems are also organized hierarchically, such that one or more high order systems control multiple low order systems, which are responsible for organizing effector output. Perceptual information flows to both low and high order control systems, and low order controllers can be capable of generating action without higher order input. Crucially, it is assumed that the architecture is the product of an evolutionary process in which higher order control has been progressively added to low order controllers, which thus have substantial preexisting control capacity. Low and high order controllers are differentially specialized: low order controllers for low order control problems, and high order controllers for high order problems. High order controllers provide flexible orchestration of low order controllers (in contrast with the subsumption model), and increased specification & refinement of low order competencies. For simple or routine activity high order controllers may be minimally active. High order controllers become maximally active in novel situations and for problems requiring complex information processing and action coordination.

The main theoretical proposal of the chapter is an account of the factors that drive the evolution of this architecture. In explaining its evolution the account also provides an explanation of many of the core functional attributes of the architecture. The account of the evolution of high order control is supported by two sources of evidence. Firstly, it will be shown that it is consistent with the general structure of the evolution of sensorimotor systems in vertebrates. Secondly, it is consistent with cognitive neuroscience findings that the prefrontal cortex exhibits hierarchical control organization and performs high level executive functions. This picture provides a framework for understanding volition. The prefrontal cortex performs integrated action control functions, and some of the properties of this control correspond reasonably well to features associated with volition. No developed theory of volition is provided here, but the account blocks the prima facie challenge presented by distributed cognition and offers a platform for further investigation of volition in terms of high order action control.

## 2 Towards a biologically-based comparative framework for cognitive architecture

If cognition is notable for being distributed, an appropriate question to ask is 'distributed compared to what?' Discussion of whether cognition is distributed or centralized needs to be placed within a conceptual framework that allows for systematic comparison. In fact the frame of reference has been largely shaped by the advocacy of rival theoretical paradigms within cognitive science: the cognitivist symbolic computation paradigm, the connectionist artificial neural network paradigm, the behavior-based robotics paradigm, the dynamical systems paradigm, and the situated cognition paradigm. Collectively, the latter four propose conceptions of cognition that are distributed in comparison with the cognitivist model. However there are significant problems with this situation. Since the units of comparison are whole paradigms the frame of reference is very coarse; the claim that cognition is distributed thus means something like 'more distributed than a von Neumann architecture' or 'more distributed than cognitivists thought it was'. This offers little basis for addressing structured questions. For instance, are there degrees of organizational distribution in functional architecture? Is it possible that differences in degree of organizational distribution are cognitively important? The cognitive processes of pencil-and-paper arithmetic show distributed organization, but are these processes *as* highly distributed as, say, the swimming of a jellyfish?

From a conceptual standpoint we need organizational concepts that allow us to specify in more precise ways the respects in which cognitive architectures can be centralized or distributed. From an empirical standpoint claims about the distribution or otherwise of cognition should be placed in a structured comparative framework. It is not difficult to find examples of cognitive processes that show some form of distributed organization, but it is less clear what the exact significance of this is. Simply collecting examples that support a rather broad hypothesis can give a misleading picture because it can overlook evidence that points in other directions. Making predictions in the context of structured evidence provides a much tougher and more informative test. In this respect the relationship between the range of actual architectures and cognitive abilities is the appropriate frame of reference for comparison. Within this framework questions such as the following arise: Is vertebrate neural architecture more or less centralized than arthropod neural architecture? Can differences in centralization between these taxa be associated with differences in behavioral abilities? A comparative framework of this kind is the bedrock on which a rigorous scientific approach to cognitive architecture should be based.

We can specify systematically the kind of evidence that should be addressed by a theory of cognitive architecture in the following way. As the highest level theory it should provide structured answers to the most fundamental questions. These include: What is cognition? What determines significant variations in cognitive ability? Which evidence is most relevant follows from the questions. In particular, the most fundamental questions correspond to the most fundamental patterns in the empirical evidence. These are of two kinds: (i) the fundamental features of sensorimotor architecture, and (ii) the empirical distribution of cognition. The central type of evidence that a theory of cognitive architecture should explain, then, is large-scale patterns in the evolution of sensorimotor organization and behavior in metazoa. Before more complicated questions about human cognitive architecture can be solved the bread-and-butter issues should be securely handled. This point is worth insisting on: if it cannot explain this kind of evidence there is reason to think that the theory doesn't have a very good grip on the nature of cognition. If it does have a good model of cognition the theory should be able to say in a reasonably precise way what it is that is under selection when cognition evolves.

When measured against these conceptual and empirical criteria the distributed cognition paradigm fares poorly. It does not provide a clear positive account of what cognition is and offers little purchase on the problem of specifying the nature of variations in cognitive ability. Consequently it doesn't provide a structured basis for explaining the empirical distribution of cognition. To be fair, distributed cognition was not framed with these questions in mind; as noted above it has rather been focused on drawing a contrast with the cognitivist paradigm. However it is legitimate assess its strength against these criteria when it is being used as a basis for inferences about cognitive architecture intended to guide further research. With respect to the topic of this volume the relevant inferences are to the effect that cognitive architecture doesn't exhibit significant hierarchy and that it doesn't feature a central system. Because distributed cognition is conceptually and empirically much weaker than has been supposed it does not provide the support for these inferences that has been commonly assumed. Moreover, as I will show below, there is substantial counterevidence.

By comparison the high order control model presented above provides a better account of the core architectural features of cognition. It associates cognition with high order control ability and so is able to provide a structured explanation of variations in cognitive ability, and of the selection pressures that impact on cognitive ability. Most importantly, it is consistent with the kind of evidence specified above, namely empirical findings concerning the core features of sensorimotor architecture and large-scale patterns in the empirical distribution of cognition

## 3 The evolution of high order control

Almost all evolutionary theories of the origins of cognition propose that it arose in response to problems of complexity (Byrne 2000, Roth and Dicke 2005). It is also common to view behavioral flexibility as the main advantage provided by cognition (Roth and Dicke 2005), although behavioral ecologists and evolutionary psychologists have claimed that cognition is primarily an aggregate of special abilities (Lockard 1971, Cosmides and Tooby 1997). The account I now present also sees the origins of cognition in problems of complexity, and identifies the major functional benefit as flexibility. But whereas most accounts focus on external complexity (environmental or social), the present account proposes a prominent role for internal functional complexity, and identifies the evolution of the fundamental mechanisms of cognition as beginning much earlier than most accounts, and in response to much more general complexity problems. Indeed, it proposes that the evolutionary process that has given rise to advanced cognition can be traced back to early metazoan evolution. Further, it proposes that the core trait under selection in the evolution of cognition is high order control capacity, rather than more specific abilities such as spatial cognition, tool use or theory of mind. Many specific abilities have played a role in the evolution of cognition, but the deepest level of organization is shaped by problems of control that are common across many abilities.

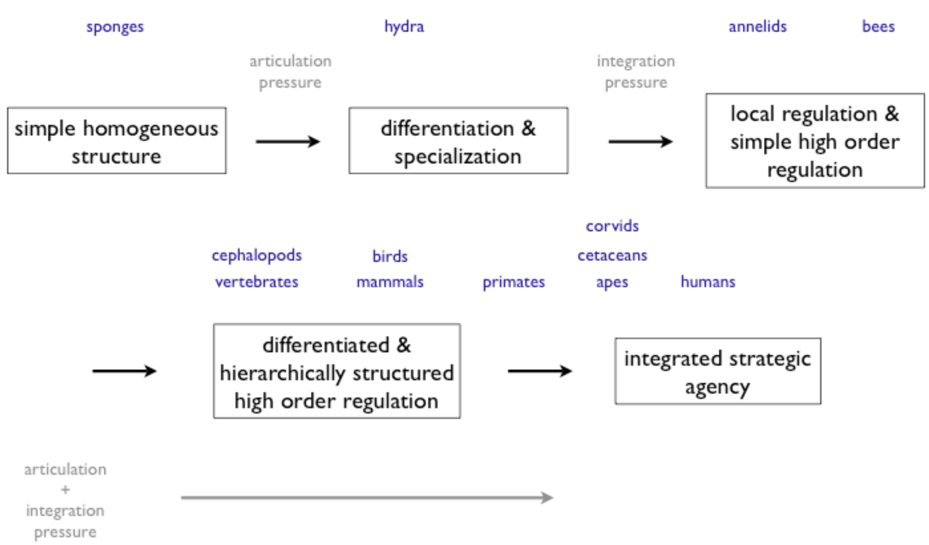


**Figure 2: Architectural transformations in the evolution of high order control.** Early multicellular animals had simple homogeneous organization. Articulation pressure drives differentiation and specialization, which creates integration pressure favoring regulative mechanisms. In vertebrates high order control becomes highly elaborated,

permitting increasingly complex & flexible strategic action control. High intelligence has evolved independently multiple times in diverse taxa. The figure depicts only one kind of evolutionary trend, so is not a *scalae natura*.

**Table 1: Major forces & constraints in the evolution of high order control**

| | |
|---|---|
| *Articulation pressure*<br>• As the conditions for successful action become more demanding there is selection for capacity to differentiate & separately control key aspects of action | *Integration pressure*<br>• As precise global state gains functional importance there is selection for mechanisms that promote coordination of collective activity |
| *Functional complexity advantage*<br>• Elaboration & specialization of action production mechanisms<br>→ Increased power, specificity, diversity, accuracy | *Functional complexity downside*<br>• Increases degrees of freedom, making global coordination a harder control problem<br>• Regulatory infrastructure is required<br>• Increased functional interdependence<br>→ Increased cost of error |
| *Behavioral & ecological factors*<br>• There are high value, difficult-to-obtain resources<br>• More complex action capacities open up new adaptive possibilities inaccessible to simpler control systems | *Variance factors*<br>• The existing architecture has a structurally available pathway for increase of control capacity |

The main architectural transitions are presented in figure 2, whilst Table 1 lists the major forces and constraints operating in the evolution of high order control. The way in which these forces drive the architectural transitions is described in the following model:

- Selection for improved action targeting creates a need for the differentiation and separate control of aspects of action production; this articulation pressure gives rise to functional complexification
- Complex functional organization offers a range of powerful adaptive benefits, including specificity, power, accuracy and diversity of action, and these benefits collectively drive continuing complexification
- Increases in complexity present high order coordination problems magnified by increased functional interdependency that increases the cost of error, resulting in pressure for integration
- Integration pressure selects for regulative mechanisms with both local and wider effects, and for integration between regulatory systems
- The hierarchical structuring of regulatory systems provides the most effective solution to the high order coordination problems presented by high complexity, and will be selectively favored as competitive pressure for increased functional capacity continues
- Selection regimes favoring high order control are likely to arise in ecological circumstances where there are high value, difficult-to-obtain resources

Note that whilst figure 2 depicts integrated strategic agency as the outcome of selection for high order control the detailed explanation for this cannot be included in this paper. Summarizing the argument, the model of the evolution of high order control serves as a framework for a more specific model of the evolution of strategic agency, which is focused on the latter stages depicted in figure 2. In essence, strategic agency is a form of high order control and evolves under the same general kinds of evolutionary pressures. High order control itself is subject to articulation pressure and becomes increasingly elaborated. Integration pressure favors the formation of integrated management of whole-system relations, including internal state and environmental interaction. The specific model will be developed in Christensen (in preparation) and is as follows:[1]

- Articulation and integration pressures acting on high order control will favor strategic action control
- Strategic action capacity is improved through the articulation of mechanisms of action-outcome control to permit relational action management
- Relational action management is subserved by relational information processing and valuation, and these capacities increase with integration capacity
- Integration pressure acting on relational action management capacity drives the evolution of capacities for high order representation of self-information, abstractive conceptual learning and executive control, and in combination these constitute the basis of integrated strategic agency

### 3.1 Defining functional complexity

The account is based on the following definitions of order and complexity:

*Order:* the scale of the correlations in a pattern; low order corresponds to local correlation and high order corresponds to wider correlation

*Complexity:* the amount of correlational structure present in a pattern

The definition of order can be understood in the following way. Low order organization corresponds to correlations that can be specified in terms of few, typically spatially local, elements of the system. High order organization corresponds to correlations that must be specified in terms of many, typically spatially widespread, elements of the system. The complexity of a pattern is determined by how much information is required to specify the pattern:

[1] Definitions used in the model: *Strategic action control:* orchestration of actions in relation to goals. *Action-outcome control:* control of action production with respect to outcomes. *Relational action management:* action control based on the relational properties of actions, entities and goals.

simple patterns can be easily described, whilst complex patterns require a lot of information.[2] Although a high order pattern is specified over many elements it need not be complex, as is the case for a simple gradient or regional difference. On the other hand complex patterns typically will be of high order, involving relations between many system elements. Some of the most important factors that contribute to the organizational complexity of a system are: (i) the number of system elements, (ii) the number of types of system elements, (iii) the number and type of interactions between system elements.

High complexity will tend to show correlational structure at multiple scales, and consequently a combination of regional heterogeneity and coherent larger scale patterning. If we restrict our focus to living systems adaptivity appears as a key constraint requiring global functional coherence. This leads to the following definition:

> *Functional complexity:* richly structured organization of functional systems and processes featuring regional heterogeneity and global coherence

### 3.2 The advantages of functional complexity

The central idea of the model outlined above is that functional complexity offers major functional advantages but carries with it a core tension that drives the evolution of increasingly complex hierarchically structured regulation. This tension stems from the fact that functional complexity involves a combination of regional diversity *coupled with* global coherence. Increases in complexity must somehow balance these two competing factors.

But why become more complex at all? It is now accepted wisdom in biology that there is no essential adaptive advantage to complexity; viruses and bacteria are 'as adaptive' as, for instance, large primates. The apparent increase in organismic complexity during the course of evolution can ostensibly be explained by the fact that, starting from a simple base, there was nowhere else to go. Simple diffusion through morphospace can produce an apparent trend towards increased complexity. However, although it is true that there is no essential direct link between complexity and adaptiveness, they may nevertheless not be wholly independent either. Being adaptively successful requires performing the right actions[3] at the right time. In a competitive environment the conditions for successful action targeting tend to become more demanding, and this can create pressure for the differentiation and separate control of key aspects of action production. For

---

[2] I am describing here in intuitive terms the definition of complexity provided by Bennett (1985). See Collier and Hooker (1999) for a more general discussion.

[3] I am using the term 'action' here broadly to mean the product of a functional process. This encompasses all of the internal functional processes of the system, such as, e.g., protein manufacture.

example, in order to pick up a glass without knocking it over it is very helpful to be able to independently control the force and direction of arm movement. The simplest action production systems in biology lack this kind of differentiated control: in the case of a paramecium the direction of swimming motion is determined randomly and the force of the motor action is fixed. Clearly, then, articulated action control can confer advantages by improving action targeting. I will refer to the relative advantage of more differentiated action production when there is selection for improved action targeting as *articulation pressure.* In these circumstances organisms with less articulated action production systems are out-competed by their more accurate conspecifics.

The effect of sustained articulation pressure is complexification. Increased complexity through differentiation and specialization permits more complex production processes through interaction between differentiated components. This allows more resources of greater variety to be brought to bear on action production. The list of adaptive benefits of complexity includes:

*Power:* the ability to concentrate energy making an action stronger, faster, more sustained, etc.

*Specificity:* the ability to produce an action type matched to a particular context

*Diversity:* a greater range of action types can be produced

*Accuracy:* further improvements to the targeting of action can be gained

These benefits are recursive inasmuch as they apply to the production mechanisms themselves. The effect of this is to facilitate further articulation as enhanced production capacity allows the manufacture of more specifically structured system components able to participate in more precisely structured functional processes. These are powerful adaptive advantages, and hence there is reason to expect selection to lead to increased functional complexity in many circumstances.

### 3.3 Functional complexity produces integration pressure, which selects for regulation

Although increasing organizational complexity can confer substantial adaptive benefits it also brings with it associated problems. The advantages of functional complexity stem from integrating heterogeneous components and processes, but diversity and coherence are in tension with one another. As the number of heterogeneous system components increases, and as the complexity of the components themselves increases, the coordination demands for achieving a globally coherent functional state increase. This is compounded by the fact that functional complexity will gain an adaptive advantage by enabling more complex morphologies (in the case of developmental mechanisms) and more complex ways of interacting with the environment (in the case of physiological and behavioral mechanisms), which will tend to expose the organism to a greater range of developmental and environmental conditions. These will require different patterns of activity at different times. The organism must consequently be able to maintain and switch between multiple functional regimes, where each regime is a particular set of coordinated functional states and processes. At the same time the cost of failing to achieve global coherence increases. This is because increases in functional complexity inherently tend to increase functional interdependency, but increased interdependency increases the likelihood that a functional failure somewhere in the system will propagate to downstream processes, resulting in a cascade of dysfunction. Thus, whilst the advantages of functional complexity depend on integration, increases in complexity make integration harder to achieve, and the costs of failing to achieve integration increase. I shall refer to the escalating need for integration as complexity increases as *integration pressure.* Achieving the benefits of increased functional complexity will be dependent on mechanisms that promote functional coherence and thereby resolve integration pressure.

There are three main ways in which coherence can be produced: through structural constraints, through parallel interactions that produce ‘emergent’ patterns, and through regulation that directly controls for a pattern. Each has strengths and weaknesses. The most straightforward way to ensure functional coherence is to limit the degrees of freedom of the system elements through structural constraints, for example structures introduced in development that constrain physiology and behavior. This has the advantage of simplifying functional processing requirements because the functional restrictions don’t need to be dynamically generated as part of ongoing functional processing. However structural constraints limit diversity, thereby inherently limiting functional complexity.[4] More complex action abilities depend on opening up degrees of freedom, and achieving coherence in these circumstances must occur, at least in part, via some means other than structural constraints.

Functional coherence can also be achieved through parallel interactions that generate ‘emergent’ outcomes (i.e., that are not directly controlled). In this case the collective organization in question is the product of many local interactions, with no functionally distinct global instruction signal. This has the advantage of imposing minimal infrastructure requirements and can take advantage of spontaneous pattern-formation processes. But whilst ‘self-organization’ is celebrated for its capacity to generate global patterns it has significant limitations as a means of resolving the problems presented by integration pressure. The most important of these are slow action and poor targeting capacity. Precisely because achieving the global state depends on propagating state changes through many local interactions the time taken to achieve the final state can be long, and

[4] Training wheels on a child’s bike are an example of a structural constraint that allows functional performance by restricting the range of available states. They also illustrate some of the limitations that can be associated with this kind of solution: once the child is able to dynamically maintain balance for herself the training wheels are a hindrance. The regulative ability to maintain balance provides a much more powerful, flexible solution to the problem of staying upright.

increases with the size of the system. Moreover, since there is no regulation of global state, the ability of the system to find the appropriate collective pattern depends on the fidelity of these interactions. Here there is a tension: if the self-organization process is robust against variations in specific conditions the process will be reliable, but it will be difficult for the system to generate multiple finely differentiated global states. Alternatively, if the dynamics are sensitive to specific conditions it will be easy for the system to generate multiple finely differentiated global states, but difficult to reliably reach a specific state. Slow action and poor targeting capacity severely limit the capacity of self-organization to achieve the kind of coherence that functional complexity requires. As described above, the adaptive advantages of functional complexity stem in large part from precise, varied interactions that may shift on rapid timescales.

Consequently the most effective means for achieving the type of global coherence required for functional complexity is through regulation, including feedback mechanisms and instructive signals operating at both local and larger scales. The key feature that distinguishes regulation from self-organization is the presence of a functionally specialized system that differentially specifies one or a restricted set of states from the range of possible states the regulated system might take, based on the sensing of system conditions and the production of control signals that induce changes in functional state.[5] Regulation can mitigate the negative effects of organizational complexity in a variety of ways. Regulatory processes can correct errors, repair damage, and adjust process activity to changing circumstances. Error correction and repair make processes more reliable, reducing the likelihood of errors that manifest as functional failures affecting other processes. The ability to adjust activity to changing circumstances can allow downstream compensation if an upstream functional failure does occur. It also permits dynamic mutual tuning of activity that can help to ensure that systems and processes remain within mutually required ranges of activity. In combination these capacities are able to provide robustness in the context of close functional linkage. This ameliorates the increasing cost of error that accompanies increasing functional complexity. A further important feature of regulation is that it is an enabler for greater levels of functional coordination, since a local regulative ability to modify functional activity in response to signals from other systems facilitates large-scale correlated functional change. Thus, selection for functional complexity will tend to give rise to derived selection for regulative ability.

However it should also be noted that the adaptivity of an organism will always be the result of all three kinds of mechanisms operating together. That is, it will be the result of some mixture of structural constraints, self-organization and regulation. I have suggested that each kind of mechanism has different adaptive tradeoffs, and such tradeoffs will play an important role in

[5] It should be noted that whilst regulation is distinguished from self-organization it can both contribute to self-organization and take advantage of it, as is explained below.

defining differentiated adaptive strategies. Thus, we can expect to find a range of adaptive strategies in biology that emphasize some kinds of coherence-inducing mechanisms rather than others. For instance *r*-selected organisms[6] will tend to rely strongly on structural constraints, in the form of simple, highly constrained morphologies with limited regulatory complexity. It is in species where selection for functional complexity has been most prominent, thereby resulting in the greatest integration pressure, that we should expect to find the most elaborated regulative systems.

### 3.4 Strong integration pressure selects for high order regulation

The simplest response to integration pressure is local regulation of a functional process, allowing it to respond to changing circumstances. Extensive selection for regulative ability is, consequently, likely to give rise to a multitude of local regulatory mechanisms, resulting in high levels of distributed control. In addition to its local functional effects this improves the capacity for globally integrated functional behavior: collectively the system can integrate functional activity through many local adjustments. Indeed, by making local systems more sensitive to ambient conditions local regulatory mechanisms can promote collective self-organization processes and help to ensure that they result in functional end-states. However this 'self-organization' mechanism has the limitations identified above: slow action and poor targeting capacity. Moreover, as the complexity of global state changes required increases it becomes increasingly inefficient because it requires local controllers to be 'too intelligent': i.e., to have sophisticated information processing capacities that are able to determine precisely what the global context is and what the appropriate local response is. Not only can this impose prohibitive demands on local information processing capacities (e.g. of individual cells), it becomes increasingly inefficient from an evolutionary point of view because, when there is heterogeneity of local control systems, it requires many coordinated adaptations in local control systems to achieve specific changes in globally coordinated behavior.

Consequently, as integration pressure increases specialized regulative systems that have wide effect will be selectively favored. By directly modulating large-scale functional activity such systems can more effectively promote globally coordinated functional activity. A specialized regulative system can function as an integration center, gathering information from wide-ranging sources and subjecting it to processing to extract highly specific control information. Functionally specialized systems can provide spatially and/or temporally and/or qualitatively precise delivery of control signals across wide regions. In addition, a regulatory system can serve as a target of selection for variations that produce globally correlated changes.

---

[6] Organisms selected for a high rate of reproduction.

At this point it is necessary to examine the definition of order given above. According to this definition the scale of a correlation determines order. Applied to control, this gives the following definition:

*Control order:* the scope of control influence (how much of the system is subject to the control signal)

Thus, a regulatory system with wide effect is a high order controller. However high order control is also used to mean metacontrol, a sense that may be more common and intuitive:

*Metacontrol:* the control of another control system

Since there will typically be widespread local regulation, and since regulatory systems with wide effect will interact with local controllers, regulative systems with wide effect will generally be high order controllers in both the control scope and metacontrol senses. Some readers may feel inclined to restrict the meaning of high order control to metacontrol. However the definition in terms of control scope is more organizationally fundamental and more important. It plays a vital role in characterizing the architecture of high order control, which cannot be understood in terms of metacontrol alone. For instance many forms of cortical control are of high order in the sense that they control for wide-ranging aspects of the animal's internal state and behavior, independently of whether this control is directly mediated by elaborated hierarchical systems.

Increases in control ability can occur through a variety of routes, including the modification or refinement of an existing control system, the expansion of a control system, the addition of a new control system, improvement of the coordination between control systems, and the hierarchical structuring of control systems. Thus, we need to ask whether there are any factors that will tend to promote hierarchical structuring in particular. As just noted, integration pressure will select for regulative systems with wide effect, and regulative systems with wide effect will interact directly or indirectly with local control systems. Hierarchical organization will tend to arise as a natural product of this. However there are in addition specific adaptive benefits provided by hierarchical organization. It is a very efficient way to increase diversity because the same components can produce different outputs as a result of differing modulatory input. Further, structuring regulatory systems hierarchically provides a way of partitioning the control problem that allows increased global coordination whilst keeping the overall management problem for any given control system tractable. Specialized higher order control systems reduce the coordination burden on local regulation. Conversely, effective local regulation reduces the problem facing high order control systems. Differential specialization between low and high order controllers allows low order controllers to optimize for local coordination problems whilst high order controllers specialize for high order coordination problems. This frees high order controllers from the problem of 'micromanaging' local responses. High order control can extend and refine existing competencies, allowing incremental, efficient improvement of functional capacity. The net effect is to allow limited capacity systems to cooperate on a complex overall problem with intimate yet partitioned structuring of control burden.

Thus, hierarchically structured regulation can provide an effective solution to the problems presented by integration pressure, thereby making available the adaptive benefits of greater levels of functional complexity. However, for high order control to be selectively favored these conditions must actually obtain within the population. Specifically: (i) The existing architecture has a structurally available pathway for the evolutionary increase of regulative capacity. (ii) The control benefits (which can include enhanced specificity, power, accuracy, diversity and coordination of action) yield overall higher returns (possibly including reduced error costs), within the niche.

Given that increased regulation can present substantial infrastructure costs as well as, potentially, negative effects of increased complexity, these conditions will by no means be universal. If the adaptive contingencies of the niche that fall within the range of variation of the population do not offer increased return for increased control ability there will be no directional selection for control ability. Indeed, there may be selection *against* high order control if the costs of increased infrastructure, energy demands and complexity are greater than the returns gained. With respect to behavior and ecology, then, two kinds of circumstances are likely to be especially important for generating selection for high order control: (i) There are high value, difficult-to-obtain resources. (ii) More complex action capacities open up new adaptive possibilities inaccessible to simpler control systems.

In addition we can expect contingency to play a major role in the evolution of high order control. The structural pathways that are evolutionarily available will be highly constrained by the nature of the existing regulatory systems. Some regulatory systems may result in evolutionary dead-ends, whilst major adaptations may depend on a prior sequence of adaptations to occur. Conversely, however, the advent of a novel regulatory system is likely to have major evolutionary effects by significantly changing the adaptive possibilities that are available.

The evolution of high order control systems is also likely to exhibit ratcheting effects. Assuming that high order control will often be selected for when there are high value benefits that are difficult to obtain, and since each additional regulatory adaptation will present costs, a regulatory adaptation increases the adaptive need to obtain high value returns, increasing the selective pressure favoring further improvements to high order control ability. The effect of such feedback can be to sustain extended directional selection, with several important evolutionary consequences. Firstly, it could act to promote episodes of rapid evolution. Secondly, however, the evolution of a major regulatory system will involve an extended suite of adaptations. Feedback that sustains directional selection can maintain selection over the extended periods of evolutionary time it takes for such adaptations to occur.

Taking these factors into consideration it should be clear that the account does not presume a *scalae natura*, orthogenesis or teleological evolution. Given the complexity of the tradeoffs involved high order control is only one of a variety of adaptive strategies for behavior control. As noted, in many cases increased high order control will not be advantageous. The selection pressures can nevertheless be seen to play a role across the full stretch of metazoan history. This is not unique case. Oriented evolution, involving consistent directional changes along one or a few dimensions, such as increase in size or speed, is a ubiquitous biological phenomenon (Simpson 1949). Intelligence is just an example of this pattern, with human intelligence constituting an extreme elaboration of a widespread adaptive strategy. Indeed, precisely because the adaptive pressures are widespread the account predicts that there should be independent evolutionary pathways in which selection for high order control has played a prominent role. This conforms to the evidence, which indicates that evolutionary increases in intelligence have occurred independently in a variety of different taxa, including birds and mammals, and cetaceans and primates (Roth and Dicke 2005, p.250).

## 4 Supportive evidence from neural evolution

Having presented the account of the evolution of high order control abstractly I now outline evidence concerning the evolution of nervous systems that supports the account. The effects of articulation and integration pressures are discernable in the earliest stages of neural evolution, in the differentiation of specialized cell types, the formation of specialized control centers and the trend towards centralized neural organization. Extended hierarchical structuring is apparent in the vertebrate autonomic and somatic motor systems.

### 4.1 Centralization is a prominent feature of early neural evolution[7]

The first multicellular animals were sponge-like creatures with little in the way of differentiated tissue structure, and behavioral abilities limited to the control of water flow through pores by adjusting the contraction of muscle cells. These cells, called myocytes, perform both the sensory and the motor functions of the organism. Nervous systems first appeared in Cnidaria, carnivorous radially or biradially symmetric animals with a sac-like body and a single body opening (the mouth) surrounded by tentacles. The evolution of neurons from ectoderm constituted a major advance in regulatory capacity by permitting the specialization of sensory function (through sensory neurons), and by permitting the rapid and precise transmission of signals to muscle cells

[7] The discussion in this section follows Swanson 2003a,b.

(through motor neurons). Sensory discrimination could become more sensitive, precise, and functionally differentiated (e.g. into different modalities). The addition of an intermediate layer of specialized communication between sensory function and motor output allows point-to-point, longer range information transmission, and creates the potential for *divergence* and *convergence* of information flow. Divergent signal paths allow a sensory neuron or sensory area to broadcast to many distant parts of the animal, permitting a rapid, coordinated whole-organism response to a sensory stimulus. Convergent signal pathways allow a given muscle cell to be sensitive to many different sensory neurons, and to the activity of other muscle cells. Thus, early neural evolution made possible much greater behavioral complexity and integration.

Cnidarian nervous systems are diffusely organized nerve nets. Flatworms represent the next grade of complexity of neural organization. They have bilateral symmetry, a dorsal-ventral (top-bottom) axis and a rostral-caudal (front-rear) axis. They move by swimming or crawling, and have a concentration of sensory neurons at the head end. The cell bodies of neurons are clustered in ganglia connected by bundles of axons called nerve cords. The clustering of neurons, in contrast with the diffuse organization of Cnidaria, is a phylogenetic trend referred to as *centralization.* The concentration of ganglia at the head end of flatworms is the simplest form of brain, and is referred to as *cephalization.* Centralization and cephalization are more pronounced in annelid worms and arthropods, and are highly elaborated in vertebrates.

These examples illustrate the principles governing the evolution of high order control described above in the following way. Conceptually, the most parallel form of organization possible is a homogeneous matrix. Sponges represent the closest biological approximation to this type of organization, and within most of the major metazoan taxa there are trends towards more complex functional forms.[8] The separation of sensory and effector functions into separate cell types, seen in the evolution of neurons, is plausibly viewed as a response to articulation pressure. This articulated control arrangement allows the activity of effector cells to be regulated in much more complex ways. In particular, it allows the activity of effector cells to be rapidly coordinated so as to achieve specific global goals. This regulative capacity fundamentally expands the functional capacities possible with a multicellular body, and is thus a keystone event in animal evolution. As such it exemplifies the kind of evolutionary impact that regulatory innovations can have.

The predacious lifestyle of Cnidaria is consistent with the hypothesis that selection for high order control is based on the capacity to obtain high value, difficult-to-obtain resources. Prey capture

[8] Whether this represents widespread selection for the adaptive benefits of functional complexity, or whether it simply represents diffusion through morphospace, can only be resolved through detailed phylogenetic analysis. The present account predicts that the former constitutes a substantial component of the phylogenetic pattern.

will typically deliver high value returns, but prey will also tend to be sporadically distributed and capable of defensive measures. The centralization of neurons into ganglia, and their rostral concentration in cephalization, observed in flatworms, annelids, arthropods and vertebrates, concentrates control and provides the basis for the formation of specialized high order regulatory systems. Centralization and cephalization are consistent with the hypothesis that the increasingly complex functional forms found in metazoan evolution have generated associated integration pressure.

### 4.2 The vertebrate autonomic system is a high order control system[9]

The autonomic system is generally thought of as an automatic, low order, non-cognitive system. Indeed, to casual observation the continuous unconscious bodily adjustments of the autonomic system might seem like a marvelous example of distributed organization. In some respects they are. But a proper appreciation of the functional organization of the autonomic system depends on the right comparative framework. When assessed in terms of anatomy and function, rather than in comparison with conscious control, the autonomic system is a centrally organized high order control system. Invertebrates lack such a specialized regulative system and have much more limited capacity for coordinated body-wide physiological changes. Thus, the autonomic system can be seen as a regulatory adaptation to the integration pressure posed by the complexity of vertebrate bodies and lifestyles.

The following description indicates some of the reactions likely to occur in response to hearing sudden loud noise behind you in a dark alley:

> In literally the time of a heartbeat or two, your physiology moves into high gear. Your heart races; your blood pressure rises. Blood vessels in muscles dilate, increasing the flow of oxygen and energy. At the same time, blood vessels in the gastrointestinal tract and skin constrict, reducing blood flow through these organs and making more blood available to be shunted to skeletal muscle. Pupils dilate, improving vision. Digestion in the gastrointestinal tract is inhibited; release of glucose from the liver is facilitated. You begin to sweat, a response serving several functions, including reducing friction between limbs and trunk, improving traction, and perhaps promoting additional dissipation of heat so muscles can work efficiently if needed for defense or running. Multiple other smooth and cardiac muscle adjustments occur automatically to increase your readiness to fight or to flee, and almost all of them are effected by the sympathetic division of the ANS [autonomic system].
>
> (Powley 2003, pp. 913-4)

[9] This section is based on Swanson 2003a, Card et al. 2003, and Powley 2003.

In broad outline the integrative action of the autonomic system is well known, but remarkable nevertheless. In light of the model of the evolution of high order control outlined above several points stand out. The level of coordination of autonomic action provides a guide to just how deeply integrated vertebrate physiology is. This is evidence for strong integration pressure. Modularity has been greatly emphasized in recent biological and psychological thinking, but although the organ systems certainly perform modularized functions such as digestion and fluid transport they also interact continuously in the production of behavior. Consequently the state of high action readiness exemplified in the ‘fight or flight’ response requires coordinated changes of activity across almost all of the physiological systems of the body. This is a good illustration of functional complexity, in which highly diverse systems are coordinated to achieve globally coherent patterns of activity. The high level of global coherence enhances the effectiveness of the adaptive response. The autonomic system also provides a clear example of the role of specialized regulatory systems in achieving high levels of global coherence. Complex, system-wide changes in activity must occur on very rapid timescales in response to specific conditions. The autonomic system provides the specialized information processing and signal delivery required to achieve this.

It also illustrates the role of hierarchical organization in enabling tractable global coordination. The most localized control provided by the autonomic system is mediated by what are known as axon reflexes: stimulation of visceral afferent neurons results in the central propagation of an action potential but it can also produce local release of neurotransmitter directly from the site of stimulation and local collaterals. These axon reflexes produce a range of inflammatory and vascular responses. The next most localized form of control is mediated by reflex arcs passing through the spinal cord. Visceral afferents project to laminae I and V of the spinal dorsal horn, sending sensory information about visceral volume, pressure, contents or nociceptive stimuli to spinal circuits that interpret the information and generate patterned responses via efferent connections back to the viscera, for example increasing heart rate and vasoconstriction. The activity of reflex arcs is integrated and coordinated by a supraspinal system known as “the central autonomic network”, consisting of a hierarchically organized network of sites in the mesencephalon, hypothalamus, amygdala, bed nucleus of the stria terminalis, septal region, hippocampus, cingulate cortex, orbital frontal cortex, and insular and rhinal cortices. Many of these centers are part of the limbic system. The integrative functions performed by this system can be divided into three types (Powley 2003, p. 928): (1) Coordination and sequencing of local reflexes, such as the autonomic responses of the mouth, stomach, intestines, and pancreas during and after a meal. (2) Integration between autonomic and somatic motor activity. For example adjusting blood flow through the body in response to postural changes to preserve blood supply to the brain. (3) Organizing autonomic activity in anticipation of key events, such as major homeostatic imbalances. These are good examples of high order control functions.

### 4.3 The vertebrate somatic motor system exhibits extensive hierarchical structuring[10]

If we assume a functional distinction between sensors, effectors and interneurons, then conceptually the most parallel form of organization approximates that of the Cnidarian nerve net: even distribution of sensory and effector cells, and diffuse spread of information by interneurons. Vertebrate sensorimotor organization is clearly nothing like this. Direct muscle-to-muscle neural connections, common in arthropods, are absent in vertebrates. Control of muscle activity is entirely located within the spinal cord and higher sites of the central nervous system. Whilst we have an intuitive sense when performing skilled activity that our body 'knows what to do' – the phenomenology that informs the lay concept of 'muscle memory' – in fact the information guiding skilled action is not stored in the muscles but in the brain stem, cerebellum, basal ganglia and cortex. In other words, skill memory is stored in high order control systems rather than distributed through the muscle system.

Vertebrate motor control shows a similar, though more elaborated, hierarchical structuring to that of the autonomic system. The nature of this hierarchy was first demonstrated in the early 20th century by experiments involving sectioning the central nervous systems of cats (Brown 1911, Sherrington 1947). When the brain stem and spinal cord is isolated from the forebrain a cat is still able to breath, swallow, stand, walk and run. However the movements are produced in a highly stereotyped, robotic fashion. The animal is not goal-directed, nor does it respond to the environment. Thus, the brain stem and spinal cord are responsible for producing basic movement coordination, but not higher-level environmental sensitivity or goal-directedness. A cat with intact basal ganglia and hypothalamus, but disconnected cortex, will move around spontaneously and avoid obstacles. It will eat and drink and display emotions such as rage. This level of motor control is thus responsible for the core elements of motivated behavior. The hypothalamus plays an especially prominent role in integrating the activity of the autonomic and somatic motor systems.

The cortex is required for the most complex forms of action control. The somatotopic organization of the primary motor cortex is well known. The proportionately much greater area devoted to the hands, face and tongue, compared with other body areas, is an anatomical correlate of the fact that the cortex plays an important role in the control of complex skilled movements. Body areas requiring fine control are represented in more detail. The cortex is connected to the spinal cord via descending pathways of several kinds. Corticospinal pathways project directly to the spinal cord, whilst rubrospinal and reticulospinal pathways project to motor centers in the brain stem that in turn project to the spinal cord.  The cortico- and rubrospinal connections

---

[10] This section is based on Gazzaniga et al. 1998, Swanson 2003a, Grillner 2003, Floeter 2003, Scheiber and Baker 2003.

transmit commands for skilled movements and corrections of motor patterns generated by the spinal cord. The reticulospinal connections activate spinal motor programs for stereotypic movements such as stepping, and are involved in the control of posture. This distinction is reflected in the two kinds of descending pathways depicted in figure 1. Functionally this allows high order control influence to be mediated either through existing motor programs (e.g. habits and reflexes) or to act more directly on the patterning of muscle action. This permits fine-grained, contextually sensitive control when necessary, and is a major source of flexibility.

Thus, the hierarchical structuring of the motor system can be understood in terms of the layering of control systems responsible for successively more complex forms of control. The relationship between the control of basic walking movement in the spinal cord and contextual and posture control in the brain stem exemplifies this. The core elements of walking motion are produced by a repeated pattern of muscle firing, and a simple circuit in the spinal cord (a ‘central pattern generator’, or CPG) can produce this basic cycle of activity. However walking must also be adapted to context and goals. Refinement of movement and posture control is provided by higher centers that integrate a wider range of sensory information and perform more complex information processing. This higher-level control acts by modulating spinal circuits, adjusting the basic walking pattern to the circumstances. But because spinal CPGs contribute a substantial component of the movement pattern higher control is relieved of this computational burden. Control responsibility is thus efficiently distributed.

Skilled and goal-directed actions present the most challenging control problems. In the case of skilled action, performance may need to be precisely adapted to the context, and require extended sequences of motor activity. Whilst much of the sequencing may be routinizable, success may still require continuous monitoring and adjustment of performance because the context for the skill may be complex and variable. Acquiring skilled action is especially challenging because it requires assembling component motor actions into larger structures that may be partly or wholly novel. Successfully achieving this will be strongly dependent on the capacity to monitor the relations between multiple actions, context and goals. Goal directed action more generally presents formidable control problems because it requires the ability to opportunistically identify action possibilities, which may shift dramatically as context varies, form instrumental goals that effectively satisfy that animal’s requirements given the contextual opportunities, and to flexibly coordinate actions in relation to those goals. Thus, effective goal-directed action may depend on complex valuation processes, high levels of bodily awareness, rich long-term memory for context, and intensive processing of episodic information.

Complex, skilled motor actions are associated with volition in motor control research. Such actions are referred to as voluntary because they are performed ‘at will’. However the concept is noted to be ambiguous, inasmuch as almost all types of motor behavior, including basic reflexes, can be influenced by will (Grillner 2003, p. 762). For example if one touches an object that unexpectedly turns out to be hot a withdrawal reflex will be triggered. However if the object is known to be hot the withdrawal reflex can be overridden if grasping it is important. We will return to these issues below. For now it is important to note that the concept of voluntary action plays an important role in motor research and is associated with anatomically distinct systems involved in the most complex forms of action control. This provides a prima facie case that volition has a grounding in motor control. Moreover it suggests that volition is a product of the evolutionary pressures for high order control that I have been characterizing.

## 5 Cognitive control and the central system

The evidence outlined so far indicates that the main features of the human sensorimotor system conform to the high order control model depicted in figure 1. This is not very surprising since the model was in part abstracted from this evidence. The more substantial proposal is the theoretical account of the properties and evolution of the architecture presented in section 3. Based on this account we can make a further prediction. High-level cognition should be an extension of the general pattern exemplified in basic sensorimotor architecture, being the product of the same articulation and integration pressures. The highest level of control should integrate the greatest amount of information, have the widest control scope and be responsible for the most complex action control problems. In other words, the high order control model predicts that there should be a central system.[11]

Support for this prediction is provided by cognitive neuroscience research on 'cognitive control'. Cognitive control research is concerned with the mechanisms of flexible, goal-directed behavior, and distinguishes these from 'automatic' forms of action production, including reflexes and habits. Cognitive control is closely associated with the prefrontal cortex, which has a pattern of broad connections with the rest of the brain that allows it to synthesize information from many sources and exert wide control influence (Miller 2000). A variety of lines of research have shown that it has the properties of a high order control system. It is, in fact, the highest order system in the brain.

---

[11] Although it is not directly implied by the account of the evolution of high order control that has been developed in this chapter, the model of the evolution of strategic agency briefly described at the start of section 3 serves as the basis for an associated hypothesis that subjective awareness is the product of mechanisms for assembling and processing high order control information. Integrated self-representation enhances strategic action capacity. We should then expect consciousness to be associated with the highest order control processes. For related proposals see Baars' 'global workspace' model of consciousness (Baars 1989), Metzinger's 'self model' theory of consciousness (Metzinger 2003, this volume) and Legrand's 'action monitoring' approach (Legrand, to appear).

### 5.1 The architecture of top-down control

Koechlin et al. (2003) experimentally demonstrated a model of the architecture of cognitive control consisting of a nested three-level control hierarchy (figure 3).

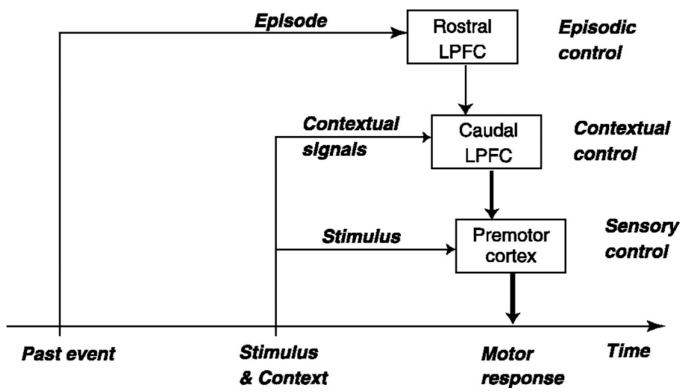


**Figure 3: Model of the architecture of cognitive control proposed by Koechlin et al. (2003, fig. 1).** Reproduced with permission. LPFC: lateral prefrontal cortex.

In the first level of control (*sensory control*) motor actions are selected in response to perceptual stimuli. This basic level of control is associated with the lateral premotor regions. In the second level of control (*contextual control*) premotor activation is regulated by information about the context. Anatomically it is associated with the caudal lateral PFC. The highest level of control (*episodic control*) regulates the contextual control system according to the temporal episode and goals. It is subserved by the rostral lateral PFC. The architecture has a cascade structure, with episodic control influencing sensory control via the contextual control level. The model was tested by presenting subjects with tasks in which stimulus, context, and episodic factors are systematically manipulated, making it possible to determine which areas of brain activation correlate with each type of factor. fMRI results showed that activation in lateral premotor, caudal lateral PFC, and rostral lateral PFC correlated with stimulus, context, and episode factors respectively, as predicted by the model.

A notable feature of this hierarchy is that higher levels perform more complex forms of integration.[12] Sensory control requires the least information, activating a stored motor program in response to an innate or learned stimulus. Contextual control is more demanding: information provided by the nature of the stimulus is not sufficient to determine what response is appropriate. It is necessary to draw on memory of the context in which the stimulus occurs. Episodic control is more complex again, requiring information both about context and current circumstances in order

[12] See Fuster 1997, 2004, Fuster et al. 2000.

to determine the appropriate response. Here we see that the same pattern found in subcortical motor organization is repeated in the cortex, consistent with the high order control model. This is counter-evidence to the distributed cognition prediction that brain organization should feature multiple control centers with no significant hierarchical organization.

### 5.2 Attentional control and fluid intelligence

One of the more significant achievements of cognitive control research is to begin to resolve the neural and cognitive mechanisms underlying fluid intelligence[13]. As recently as 1997 Deary and Caryl were able to write in a review that there was 'a dearth of explanatory accounts to link cognitive performance differences with variance in brain mechanisms' (1997, p. 365). Hypotheses under consideration included faster neural conduction, reliability of neural conduction, 'neural adaptability', and greater metabolic efficiency. Duncan et al. (2000) substantially narrowed the field of possibilities by showing that tasks that impose high demands on fluid intelligence produced focal activation in the lateral frontal cortices. Given the general association between the lateral PFC and cognitive control this implicated cognitive control in intelligence. Providing a more specific cognitive association, Gray et al. (2003) used an individual differences approach to show a relationship between fluid intelligence and attentional control.

This evidence is in many respects remarkable. That a phenomenon as complex as intelligence should be closely associated with a cognitive function as apparently straightforward as attentional control is surprising. Yet attentional control is necessary for a variety of functions that are required for complex cognitive processes. For example, in order to learn difficult concepts and form complex plans it is necessary to notice relationships between disparate locations, entities and events across extended periods of time. This requires close attention to specific features based on background knowledge and expectations.

Once again, this evidence supports the high order control model. The high order control model explains major differences in intelligence between species in terms of differences in the elaboration of high order control. Evidence that I am unable to discuss at length in this chapter indicates that the anterior cingulate cortex has been under selection in the evolution of great apes (Allman et al. 2001), and is associated with generalized high order control functions, in particular conflict monitoring (Botvinick et al. 2004). This supports the hypothesis that the evolution of intelligence in great apes has been associated with selection for high order control functions. The findings of Gray et al. (2003) indicate that high order control capacity is also able to account for

[13] Defined in terms of reasoning and flexible problem solving ability (Cattell 1971).

individual differences in intelligence in humans. Thus, the high order control model provides an empirically supported explanation of variations in cognitive ability.

### 5.3 Abstractive task learning

Many studies have shown that the PFC functions to assemble task-relevant information. Miller (2000) suggested the PFC extracts the features of experience relevant for achieving goals, detecting the relationships important for performing the task (the 'task contingencies'), and forming abstract rules, such as 'do not pick up the telephone in someone else's house'. The 'adaptive coding' model of Duncan (2001) provides some information concerning the neural basis of task-oriented processing in the PFC. According to the adaptive coding model the response properties of single neurons are highly adaptable throughout much of the PFC, and become tuned to information relevant to the task. Duncan notes that the model has connections to the idea that the PFC acts as a kind of 'global workspace' (Baars 1989; Dehaene et al. 1998).

### 5.4 In sum

Taken together these lines of research support the view that the PFC is a highest order executive control system with the kinds of properties that would be expected on the basis of the model of the evolution of high order control described in section 3. To reiterate, the highest level of control should integrate the greatest amount of information, have the widest control scope and be responsible for the most complex action control problems. This conclusion is at odds with current distributed cognition principles. In light of the theoretical considerations outlined in section 3, and the empirical evidence presented in sections 4 and 5, there is a basis for suggesting that these principles should be reexamined. A reasonable conclusion to draw from the evidence is that the major claims of no significant hierarchical structure and no central system are falsified. There is certainly no central system of the kind envisioned by classical cognitive science, but a central system nevertheless. Yet in other respects the account developed here can be seen as supporting and extending the basic distributed cognition approach.

## 6 Volition

This chapter has explored terrain not often considered in discussions of volition. But volition concerns action control, and hence a proper understanding of it depends on placing it within the context of the evolution of action control mechanisms. The account presented above provides an evolutionary and cognitive framework within which volition can be understood as a natural phenomenon. We can divide this project into two aspects: understanding the nature and status of the folk concept, and developing theoretical accounts of volition.

Developing an adequate understanding of structure of the folk concept of volition will depend on empirical concept research, but in advance of such research we can identify broad characteristics of the concept that such research will plausibly support. Features commonly associated with volition include intention, desire, choice, evaluation, command, resolution, effort, strategic flexibility, conscious awareness, and responsibility. Furthermore these features are schematically organized, such that paradigm cases of volitional action control exhibit the properties of volition in an integrated fashion. For instance the agent may form a goal that is supposed to bring about some condition that is valuable to the agent. The agent will evaluate the goal for efficacy and monitor the process of performing the action. In difficult conditions the agent may show resolve by increasing focus and strengthening goal-oriented action control processes. The agent might show strategic flexibility by switching to an alternate approach to achieving the goal. And so on. As well as schematic structure the concept is likely to have prototype structure: some examples will be better than others, and there will be marginal cases that are difficult to decide. The important point here is that to a first approximation these properties correspond reasonably well to properties of action control processes mediated by the prefrontal cortex. This suggests that scientific research is likely to result in a conservative naturalization of volition.[14] Attention and memory provide relevant comparisons here: these are folk concepts that have gained a secure scientific footing as major scientific research fields. This research has revealed a great deal of structure not present in the folk concepts but it has not shown them to be basically wrong: attention really does involve selective focus, and memory really does involve retrieving information about the past. Plausibly, volition really does involve forming goals and controlling action in relation to intentions.

With respect to developing a theoretical understanding of volition, the high order control account can help cast light on some of its more puzzling aspects. One of these is the so-called ‘freedom’ of the will. This issue is usually posed in terms of freedom with respect to the fundamental laws of nature, but if we set that aside we can identify a more adaptively relevant form of freedom: flexibility of action control. Recall that the concept of voluntary motor action used in motor control research is defined as action that can be performed ‘at will’. The high order control architecture is organized to open up action control flexibility, and this is exemplified in the

---

[14] It is important to differentiate the folk concept from the complex, theory-laden associations that volition has in philosophy. Philosophical theorizing is likely to generate much stronger and more structured commitments (e.g. to metaphysical claims) than folk concepts. Given that philosophical theorizing about volition has occurred prior to any significant scientific information about the subject, it is a sensible naturalist stance to treat it with suspicion pending a systematic reevaluation.

somatic motor hierarchy. Low levels provide stereotypic forms of motor activation such as basic walking movement, whilst higher levels adjust the action to the circumstances (e.g. brain stem postural control) and set goals such as direction and speed (determined by the cortex). The model of Koechlin et al. (2003) shows that this hierarchical organization is extended in the frontal cortex, with increasingly complex, flexible forms of control being performed by successively anterior lateral regions moving forward from the primary motor cortex. The rostral lateral PFC performs ‘episodic control’, adjusting goal-directed action in relation to local contingencies. Research such as that of Duncan (2001) shows that this control is based on rich representations of task-related information. From an adaptive perspective the highest level of action control should be extremely flexible because task contingencies can be very complex and can change dramatically. The highest-level system should be able to form and reform goals for action based on shifts in any of a large range of agent-based and environment-based factors. The formation of ‘volitions’, then, should be based on the agent’s goals, values and the environmental context but shouldn’t be consistently determined by any particular factor or type of factor. Thus, action performed ‘at will’ is determined episodically in relation to a constellation of factors, and so can exhibit high levels of spontaneity and variability.

However the high order control account presented here does not by itself furnish a theory of volition. Such a theory must address many further issues, including the subjective perception of volition (Metzinger, this volume; Wegner, this volume), social and psychological processes of self formation (Ross, this volume), and the ability to pursue longer-term objectives that require resisting short-term temptations (Ainslie, this volume). Given the diversity of perspectives on volition found in this volume there is clearly a long way to go before a theoretical synthesis is likely to emerge. The high order control proposal doesn’t grapple with some of the more complex phenomena associated with volition, but I suggest that it is an important element in the mix and can provide a basis for integrating a range of agency-related phenomena within an adaptive, control-based framework.

## Acknowledgements

Thanks to David Spurrett for a number of very helpful editorial suggestions, and collectively to the organizers of the second Mind And World conference for an absorbing and enjoyable event.

## References

Allman, J. M., Hakeem, A., Erwin, J. M., Nimchinsky, E., & Hof, P. (2001). The Anterior Cingulate Cortex: The Evolution of an Interface between Emotion and Cognition. *Annals of the New York Academy of Sciences, 935*, 107-117.

Baars, B. J. (1989). *A Cognitive Theory of Consciousness*. New York: Cambridge University Press.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford Univ. Press.

Bennett, C. H. (1985). Dissipation, information, computational complexity and the definition of organization. In D. Pines (Ed.), *Emerging Syntheses In Science. Proceedings of the Founding Workshops of the Santa Fe Institute.* Redwood City, Calif.: Addison West Publishing Company.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences, 8*(12), 539-546.

Brooks, R. A. (1991). Intelligence Without Reason. In *Proceedings of the Twelfth International Conference on Artificial Intelligence* (pp. 569-595): Morgan Kauffman, San Mateo.

Brown, T. S. (1911). The intrinsic factors in the act of progression in the mammal. *Proc. Royal Soc. London, Series B, 84*, 308-319.

Byrne, R. W. (2000). Evolution of primate cognition. *Cognitive Science, 24*(3), 543-570.

Cajal, S. R. (1894). *Les Nouvelles idées sur la structure du système nerveux chez l'homme et chez les vertebrés*. Paris: Reinwald.

Card, J. P., Swanson, L. W., & Moore, R. Y. (2003). The hypothalamus: an overview of regulatory systems. In L. R. Squire, F. E. Bloom, S. K. McConnell, J. L. Roberts, N. C. Spitzer & M. J. Zigmond (Eds.), *Fundamental Neuroscience* (2nd ed., pp. 897-908): Elsevier Academic Press.

Cattell, R. B. (1971). *Abilities: Their Structure, Growth and Action*. Boston: Houghton Mifflin.

Christensen, W. D. (in preparation). The evolution of strategic agency.

Clark, A. (1997). *Being there: Putting brain, body and world together again*. Cambridge, MA: MIT Press.

Collier, J. D., & Hooker, C. A. (1999). Complexly organised dynamical systems. *Open Systems and Information Dynamics, 6*, 241-302.

Cosmides, L., & Tooby, J. (1997). The modular nature of human intelligence. In A. B. Scheibel & J. W. Schopf (Eds.), *The Origin and Evolution of Intelligence* (pp. 71-101). London: Jones and Bartlett.

Deary, I. J., & Caryl, P. G. (1997). Neuroscience and human intelligence differences. *Trends in Neurosciences, 20*(8), 365-371.

Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl. Acad. Sci. U. S. A., 95*, 14529–14534.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience, 2*, 820-829.

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., et al. (2000). A Neural Basis for General Intelligence. *Science, 289*(5478), 457-460.

Floeter, M. K. (2003). The Spinal Cord, Muscle, and Locomotion. In L. R. Squire, F. E. Bloom, S. K. McConnell, J. L. Roberts, N. C. Spitzer & M. J. Zigmond (Eds.), *Fundamental Neuroscience* (2nd ed., pp. 767-790): Elsevier Academic Press.

Fuster, J. M. (1997). *The Prefrontal Cortex*: Lippincott-Raven.

Fuster, J. M. (2004). Upper processing stages of the perception-action cycle. *Trends in Cognitive Sciences, 8*(4), 143-145.

Fuster, J. M., Bodner, M., & Kroger, J. K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature, 405*, 347-351.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (1998). *Cognitive Neuroscience: the biology of the mind*. New York: W. W. Norton & Company.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*(3), 316-322.

Grillner, S. (2003). Fundamentals of motor systems. In L. R. Squire, F. E. Bloom, S. K. McConnell, J. L. Roberts, N. C. Spitzer & M. J. Zigmond (Eds.), *Fundamental Neuroscience* (2nd ed., pp. 755-765): Elsevier Academic Press.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science, 302*(5648), 1181-1185.

Legrand, D. (to appear). The bodily self: The sensori-motor roots of pre-reflexive self-consciousness. *Phenomenology and the Cognitive Sciences*.

Lockard, R. B. (1971). Reflections on the fall of psychology: Is there a message for us all? *Am. Psychol., 26*, 168-179.

Metzinger, T. (2003). *Being No One: The Self-Model Theory of Subjectivity*: MIT Press.

Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience, 1*, 59-65.

Norman, D., & Shallice, T. (1986). Attention to action: willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and Self-Regulation* (Vol. 4, pp. 1-18). New York: Plenum Press.

Posner, M. I., & Raichle, M. E. (1994). *Images of Mind*. New York: Scientific American Library.

Powley, T. L. (2003). Central control of autonomic functions: organization of the autonomic nervous system. In L. R. Squire, F. E. Bloom, S. K. McConnell, J. L. Roberts, N. C. Spitzer & M. J. Zigmond (Eds.), *Fundamental Neuroscience* (2nd ed., pp. 913-932): Elsevier Academic Press.

Roth, G., & Dicke, U. (2005). Evolution of the brain and intelligence. *Trends in Cognitive Sciences, 9*(5), 250-257.

Scheiber, M. H., & Baker, J. F. (2003). Descending control of movement. In L. R. Squire, F. E. Bloom, S. K. McConnell, J. L. Roberts, N. C. Spitzer & M. J. Zigmond (Eds.), *Fundamental Neuroscience* (2nd ed., pp. 792-814): Elsevier Academic Press.

Sherrington, C. S. (1947). *The Integrative Action Of The Nervous System* (2nd ed.). New Haven: Yale University Press.

Simpson, G. G. (1949). *The Meaning of Evolution*. New Haven: Yale University Press.

Swanson, L. W. (2003a). *Brain Architecture - Understanding the Basic Plan*: Oxford University Press.

Swanson, L. W. (2003b). The architecture of nervous systems. In L. R. Squire, F. E. Bloom, S. K. McConnell, J. L. Roberts, N. C. Spitzer & M. J. Zigmond (Eds.), *Fundamental Neuroscience* (2nd ed., pp. 15-44): Elsevier Academic Press.