Article:

Prescott, T. J., Redgrave, P. and Gurney, K. orcid.org/0000-0003-4771-728X (1999) Layered control architectures in robots and vertebrates. Adaptive Behavior, 7 (1). pp. 99-127. ISSN 1059-7123

https://doi.org/10.1177/105971239900700105

Layered Control Architectures in Robots and Vertebrates

Tony J. Prescott, Peter Redgrave, and Kevin Gurney

University of Sheffield, UK.

Abstract

We review recent research in robotics, neuroscience, evolutionary neurobiology, and ethology with the aim of highlighting some points of agreement and convergence. Specifically, we compare Brooks' (1986) subsumption architecture for robot control with research in neuroscience demonstrating layered control systems in vertebrate brains, and with research in ethology that emphasizes the decomposition of control into multiple, intertwined behavior systems. From this perspective we then describe interesting parallels between the subsumption architecture and the natural layered behavior system that determines defense reactions in the rat. We then consider the action selection problem for robots and vertebrates and argue that, in addition to subsumption-like conflict resolution mechanisms, the vertebrate nervous system employs specialized selection mechanisms located in a group of central brain structures termed the basal ganglia. We suggest that similar specialized switching mechanisms might be employed in layered robot control architectures to provide effective and flexible action selection.

keywords: subsumption architecture, brain evolution, behavior systems, defense system, action selection, basal ganglia.

citation: Prescott, T.J., Redgrave, P., & Gurney, K. (1999). Layered control architectures in robots and vertebrates, *Adaptive Behavior*, **7**, 99-127.

Layered Control Architectures in Robots and Vertebrates

The field of adaptive behavior seeks a convergence of ideas from the different disciplines that study artificial and natural autonomous systems. Demonstrating convergence allows the interchange of concepts and ideas and enriches our understanding of both the biological and the synthetic (Arbib, 1989; Meyer & Guillot, 1990). Within this tradition the present article reviews research in robotics, neuroscience, evolutionary neurobiology, and ethology, with the aim of highlighting some key areas of agreement, and argues that this cross-disciplinary perspective could help to resolve some of the current dilemmas facing research in autonomous robotics.

Rodney Brooks' (1986, 1989, 1990, 1991ab, 1995) work in engineering robot "creatures" needs little introduction to researchers in adaptive behavior. In the mid-eighties Brooks introduced a new methodology—based on an analogy with natural evolution—for building "self-sustaining" mobile robots that operate in real-time and in un-customized human environments. This research has had enormous influence in robotics and, together with other contemporary work that proposed a move towards more distributed and situated systems (e.g. Braitenberg, 1986; Minsky, 1986), has inspired a new research paradigm in artificial intelligence (see e.g. Meyer & Guillot, 1990; Maes, 1992). A key contribution of Brooks' work is his proposal for a layered, distributed control architecture for mobile robots, termed the *subsumption architecture (SA)*. Section 1 of this article briefly outlines the key features of the SA.

A substantial body of the neuroscience literature can be interpreted as demonstrating layered control systems in the vertebrate brain. In many ways the notion of layering is a common, often unspoken, assumption in contemporary neuroscience, however, the implications of the layered nature of the brain are not always acknowledged in a field dominated by the study of the mammalian cortex. Section 2 considers work that follows in the tradition of John Hughlings Jackson (1884/1958), a neuropsychologist who is particularly associated with the notion of layered competence, while Section 3 looks for similarities between the robot design process proposed by Brooks and the evolutionary history of the vertebrate brain.

An understanding of adaptive behavior is central to the *behavior systems* approach which stems from pioneering work in ethology by Lorenz, Tinbergen, and Baerends (see Baerends, 1976), and has been influential in some recent research in psychology and neuroscience (see Timberlake, 1993). A key principle is that the functional organization of the vertebrate brain can be decomposed into multiple, semi-independent, systems dedicated to major biological functions such as feeding, reproduction, defense, and body care. Section 4 gives a brief outline of the behavior systems approach relating it to Brooks' proposal for behavior-based robot control.

A general thesis is best served by a specific example. In Section 5 we argue that the layered neural architecture that implements the defense behavior system in the rat bears many interesting resemblances to Brooks' SA. A number of specific correspondences are outlined in detail.

Finally, in Section 6, we consider the action selection problem for both natural and artificial control systems, distinguishing between emergent and specialized action selection, and between distributed and centralized selection mechanisms. We then note that the *basal ganglia* (BG), a group of functionally-related, central brain structures, have a controlling influence on neural systems at multiple levels of the vertebrate nervous system and so form an important exception to the overall vertical decomposition of the brain. The BG, we propose, act as a specialized action selection device that provides flexible conflict resolution between functional units that are widely distributed in the brain. Although providing a centralized selection system within the global brain architecture, the BG exploits the advantages of distributed switching at a local level. Understanding the function of the BG within the vertebrate brain, we suggest, could help in the design of effective action selection mechanisms for robots, supplementing the use of layered, subsumption-style control.

1 BROOKS' SUBSUMPTION ARCHITECTURE

Brooks (1986) introduced the subsumption architecture (SA) as the central element of his proposal for building "complete creatures" capable of sustained activity in everyday human environments. With only limited modification (Brooks, 1989) the principles of the SA have since been employed in the design and control of a large number of mobile robots (see Brooks, 1990, 1991b) and have been widely copied. The key aspects of the SA, of most relevance here, are as follows:

Distributed, layered control. Control in a Brooks' robot is distributed across several layer*s* each composed of multiple modules often mounted on different processors. Layers operate in parallel and asynchronously. Within a layer there is no central control module.

Behavioral decomposition. The different layers of the control system are designed to support different "task-achieving" behaviors (such as obstacle avoidance, wandering, exploring, map-building); the problem of controlling the robot is thus decomposed into behavioral units rather than into different "functional"[1] units (such as perception, modeling, planning, and motor control). Within a layer there may be a more traditional decomposition; for instance, into sensor and actuator components. However, different layers will use different decompositions based on specialized sub-sets of the available sensorimotor apparatus.

Increasing "levels of competence." Each ascending level of the control system adds to the behavioral capabilities of the robot resulting in a higher level of overall competence. Damage or failure at a higher level reduces the robot to functioning at the level of the next highest layer. The higher layers of the SA often operate by modulating the activity of lower layers—hence their contact with the motor resources can be relatively indirect. Since higher level modules can implicitly rely on the operation of lower-level behavioral primitives, they can be designed to generate more complex or subtle motor acts.

Incremental construction. A key constraint on the design process is that, as each additional level of competence is incorporated, the total system should be "a strict augmentation of the previous one" (Brooks, 1989 p. 253). Designing the control system is therefore an incremental process in which each intermediate architecture is extensively tested and debugged before the next layer is added.

Conflict resolution and communication between levels by subsumption mechanisms. Higher layers of the control system can *subsume* the roles of lower ones by suppressing their

---

[1] Brooks' distinction between functional and behavioral components is not in common use in biology. Here the term function will generally be used to indicate the purpose or use of a mechanism as opposed to its form or structure. Where the distinction that Brooks has proposed is to be considered this will be made clear in the text.

outputs and (optionally) substituting their own. Each lower level continues to function as higher levels are added, "unaware" that those above may be interfering with its data paths.

Little sensor fusion and no central models. Brooks' (1986) places less emphasis on the combination of multiple sensor signals to determine the most accurate estimate possible of world state (*sensor fusion*) than on the independent use of different sources of sense data to provide robustness to changing environmental conditions, sensor noise, or hardware failure (see also Prescott, 1996). One consequence of this view is the principle that the robot should have no need for central world models into which all available sense data is compiled. Rather than exploiting shared representations, behaviors at different levels are separated by "abstraction barriers" (Brooks, 1990), unable to influence each other's internal workings by anything more than simple subsumption mechanisms. Internal state at each higher level cannot be accessed by lower layers, although higher layers can access the data paths of those below.

The Subsumption Metaphor in Biology

Though the main impact of Brooks' work has clearly been in robotics and AI, it has also had a significant influence on the study of natural intelligence. Work in this wide area that acknowledges the influence of Brooks' approach includes studies of human perception and motor control (Ballard, Hayhoe, & Whitehead, 1992; King, Dykeman, Redgrave et al., 1992), human development (Rutkowska, 1994), and research in computational neuroethology (e.g. Altman & Kien, 1989; Arbib & Liaw, 1995; Cliff, 1991; Franceschini, 1992).

Brooks himself does not make strong claims for the SA as a model for understanding natural autonomous systems. Indeed, he explicitly states that, although the SA draws on an evolutionary metaphor, it is not a biological model. He also warns of the dangers of treating biological intelligence as a lodestar for AI (Brooks, 1995). However, Brooks also insists that his interest is in general intelligence (Brooks, 1995), that he sees animal intelligence as an important "existence proof of the possibility of intelligent entities" (Brooks, 1990, p. 5), and that we should expect to gain insights for robot design by studying the nervous systems and behavior of animals (Brooks, 1991a; Brooks, 1995). The search for further links between Brooks' robot architectures and our understanding of animal intelligence therefore fits naturally with the situated robotics approach.

The SA is not unique, of course, in being a layered robot control architecture. For example, Arbib and Liaw (1995), Roitblat (1991), McGonigle (1990), and Albus (1991), have each proposed architectures for robot control inspired by biological intelligence and sharing a number of interesting similarities with the architecture of the vertebrate brain. Several architectures have also been proposed that are principally refinements or extensions of the SA (e.g. Rosenblatt & Payton, 1989; Connell, 1990). This article focuses on the original SA, however, as it is probably the best known and most imitated architecture in behavior-based robotics. Drawing comparisons with such a widely understood model will, we hope, encourage robot designers to look with greater interest at the organization of the vertebrate nervous system as a source of inspiration for the design of robot control architectures.

2 THE VERTEBRATE BRAIN VIEWED AS A LAYERED ARCHITECTURE

The Jacksonian Perspective in Neuroscience

In 1884, in a famous lecture on the "evolution and dissolution of the nervous system" the neurologist John Hughlings Jackson (1884/1958) proposed a layered view of the nervous system, in which the brain is seen as implementing multiple levels of sensorimotor competence. Jackson’s view, inspired by the Darwinian revolution in nineteenth century science,[2] was based not on the usual morphological divisions but on functional grounds, “as to the degree of indirectness with which each [division] represents the body, or part of it” (p. 53). He divided the nervous system into lower, middle, and higher centers, and proposed that this sequence represented a progression from the “most organized” (most fixed) to the “least organized” (most modifiable), from the “most automatic” to the “least automatic,” and from the most “perfectly reflex” to the least “perfectly reflex.” This progression sees an increase in competence in a manner that we might now understand as a behavioral decomposition—higher centers are concerned with same sort of sensorimotor coordinations as those below, though in a more indirect fashion:

> That the middle motor centers represent over again what all the lowest motor centers have represented, will be disputed by few. I go further, and say that the highest motor centers (frontal lobes) represent over again, in more complex combinations, what the middle motor centers represent. In recapitulation, there is increasing complexity, or greater intricacy of representation, so that ultimately the highest motor centers represent, or, in other words, coordinate, movements of all parts of the body in the most special and complex combinations. (Jackson 1884/1958 p. 53)

Jackson viewed the evolution of the nervous system as an incremental process in which lower levels are retained intact but are suppressed by higher systems. Within the different centers Jackson further considered there to be functionally distinct layers. He argued for a dissociation of higher layers from those below such that a breakdown at a higher layer—a “dissolution” in Jackson’s terminology—caused a reversion to the next highest layer of control.

There are further important parallels between Jackson’s writing and contemporary approaches in robotics. For instance, he was an early of advocate of the notion of distributed representation. Overall, his writings show a conviction that “higher” thought is grounded in perception and action—a perspective which, while radical for his era,[3] is clearly in sympathy with recent proponents of situated action (e.g. Brooks, 1990, 1995; Chapman, 1991):

> A man physically regarded is a sensorimotor mechanism. I particularly wish to insist that the highest centers—physical basis of mind or consciousness—have this kind of constitution, that they represent innumerable different impressions and movements of all parts of the body [..] It may be rejoined that the highest centers are “for mind.” I assert that they are “for body,” too. If the doctrine of evolution be true, all nervous centers must be of sensorimotor constitution. (Jackson, 1884/1958, p. 63)

Jackson’s views on the functional organization of the nervous system continue to influence and inspire neuroscientific research (see e.g. Teitelbaum, Schallert & Whishaw, 1983; Rudy, Stadler-Morris & Albert, 1987; King et al., 1992; Berntson, Boysen & Cacioppo, 1993), and there is now a mass of empirical evidence—anatomical, physiological and

---

[2] A number of Hughlings Jackson's contemporaries held similar views on brain organization, for review see Magoun (1958), Berntson, Boysen and Cacioppo (1993).

[3] The prevailing view was that the brainstem and spinal cord controlled motor functions whilst the cerebral cortex was reserved for higher cognitive functions (Jackson, 1884/1958; Humphrey, 1986).

behavioral—that supports the notion of layered control systems in the vertebrate brain. Some of the experimental support for this view is briefly summarized below.

Dissociations

The Jacksonian view predicts a dissociation between higher and lower level components of a system where the lower level competence is left intact by damage at a higher level. Discoveries of such dissociations were amongst the earliest findings of neuroscience and are now understood to abound in the vertebrate nervous system (for reviews see Gallistel, 1980; Wishaw, 1990). For instance, it has been demonstrated that removing the cerebral cortex from a cat or a rat—thereby eliminating many major sensory, motor and cognitive centers—leaves intact the ability to generate motivated behavior. The animal still searches for food, eats to maintain body weight, shivers when cold, fights or escapes when attacked, and so on. These behaviors appear awkward and clumsy when compared to controls and are often poorly adjusted to circumstances. However, the ability to generate appropriate, motivated action sequences is retained. When most of the forebrain is removed, complete behaviors can no longer be produced, but the capacity for individual actions such as standing, walking, grooming, and eating is spared. With all but the hindbrain and spinal cord removed, the animal cannot coordinate the movements required for these actions—for instance, it cannot stand or walk unaided—however, most of the component movements that make up the actions are still possible. Similar results have been observed in several classes of vertebrates (see, for instance, Overmier & Hollis, 1990, for a discussion of related findings in fish) leading to the widely accepted conclusion that the anatomically lower levels of the vertebrate nervous system organize simpler movements, while higher levels impose more complex forms of behavioral control.

Jacksonian Progressions

A further implication of a Jacksonian view is that during ontogeny (development) the brain matures through the sequential addition of higher centers (Teitelbaum et al., 1983; Rudy et al,. 1987). This form of development has been observed post-natally in the maturation of rats and rhesus monkeys (see Rudy et al, 1987 for review). For instance, in rats, reflexive responses to stimuli (visual, auditory, or gustatory) have been shown to mature several days before the same stimuli are able to mediate learned behavioral reactions. Unlearned reflexive responses can be generated by the brainstem components of sensory systems, whereas learned behaviors of this kind generally require higher-level components. Evidence for a comparable developmental sequence can be found in the rat's ability to negotiate the Morris water maze task (Rudy et al., 1987). Specifically, infant rats can learn to approach a visible goal before they are able to use distal visual cues to locate and approach an invisible goal, the former ability can be mediated by low-level, associative learning mechanisms, while the latter ability requires the operation of at least one higher level structure (the hippocampus).

It is important to note that the hypothesis that the brain is a layered architecture is not identical to the conjecture that brain evolution has followed an incremental path, even though many, including Jackson, have seen it in that light. The architecture of the brain may be layered simply because this provides an efficient and robust design. Likewise, a developmental progression is unlikely to be a strict recapitulation of the phylogenetic origins of the brain (Deacon, 1990; Butler & Hodos, 1996) but will be the consequence of a combination of constraints on the ontogenetic process, including, perhaps, the need for brain systems to "boot-up" in an appropriate order. We will consider the evidence for incremental change in the nervous system separately when we address the topic of brain evolution in Section 3.

Hierarchies, Heterarchies, or Layers

Hughlings Jackson, like many others since, suggested that the layered architecture of the brain provides a hierarchical organization of neural function. The many advantages of hierarchical decomposition of control have been summarized by Szentágothai and Arbib (1975) (see also Arbib, 1989). Gallistel (1980), reviewing some of the classical neuroscience literature, concludes that the nervous system can be viewed as a "lattice hierarchy" in which lower-level units are recruited by multiple different, and potentially competing, higher level masters. According to this view, the activation of lower level systems occurs through a combination of activating inputs at the same level (some of which may be of sensory origin) and potentiating signals from higher levels that can allow or disallow the units' response. Roitblat (1991) has suggested that Gallistel's lattice design could make an effective robot control architecture.

Opposition to the view of the brain as a hierarchical control system has focused on the nature and extent of feedback loops between lower and higher level neural systems. The existence of such loops has been interpreted as evidence for heterarchical rather than hierarchical control (see e.g. Kien, 1986; Cohen, 1992). Szentágothai and Arbib (1975), who are otherwise favorably disposed toward hierarchical schemes, note that such loops certainly make it more difficult to decide exactly "who is commanding whom" in the system. We conclude that a purely hierarchical arrangement does not accurately describe the vertebrate nervous system, which may be best viewed as a sophisticated trade-off exploiting useful characteristics of both hierarchical and heterarchical schemes. In this article we therefore use the more neutral term “layered”, which, following Jackson and Brooks, we understand to mean the decomposition of a control system into multiple levels of competences with lower levels dissociable from those above.

Our definition of a layered control system specifically allows for feedback loops that support two-way interactions between different levels. It is interesting to ask whether there is anything equivalent to this in Brooks' SA in which the emphasis, as in many models of brain architecture, is on one-way links from higher to lower levels. First, we note that the SA allows modules in each higher layer to inspect the data paths of modules in the layer below. This mechanism could allow the higher level system to respond to changes it has itself initiated in the lower level system. Second, an important source of between level feedback in Brooks' architecture arises from the notion of interaction through the world. Modules, that are unconnected within the control system, may yet influence each other through the effects of their actions on the state of the agent or the environment. This possibility, which is often overlooked when the nervous system is viewed as a control hierarchy, provides an alternative means of completing feedback loops between higher and lower level systems.

Further Dimensions of Nervous System Organization

The nervous system is an intricately complex three-dimensional structure. Although we have emphasized the layered nature of the brain we recognize that there will be other important governing principles in brain organization.

Following the embryologist Karl Ernst von Baer a number of neuroscientists have recognized a concentric dimension to brain organization (see Magoun, 1958; Berntson et al., 1993). Von Baer noted that brain development proceeds from the center outwards, and that as neurons migrate to lateral and peripheral positions they become differentiated and specialized. This suggests an organizing principle according to which more generalized systems are placed centrally and more specialized ones at the periphery.

While a concentric dimension can be seen as complimentary to a vertical (layered) organization, the neurologist Wilder Penfield (1958) proposed an organizing principle that is less easily accommodated within a Jacksonian view. Penfield suggested that a group of central, sub-cortical brain structures serves to coordinate and integrate the activity of both higher- (cortical) and lower-level neural systems. This notion is captured in the proposal of a *centrencephalic* dimension to nervous system organization (see also Berntson et al., 1993; Thompson 1993). Penfield's theory has been further developed by Thompson (1993), who identifies the centrencephalic "core" with brain nuclei belonging to, or associated with, the basal ganglia. In Section 6 we will suggest that many of these structures can be accommodated with the view of the brain as a layered architecture by assigning them a key role in action selection.

3 THE EVOLUTION OF THE VERTEBRATE BRAIN

The above discussion suggests that a major organizing principle in the vertebrate brain is a vertical decomposition into layered sub-systems. We suggest that this layered architecture has interesting parallels with Brooks' subsumption architecture for robot control. A more detailed exposition of this position for the specific case of the rat defense system will be presented in Section 5.

In addition to suggesting how a robot control system might be organized, Brooks' SA contains a proposal for the design process for such systems. A key idea of the SA—which draws on an analogy with natural evolution—is that a complex control system can be constructed by progressively incrementing an initially simple system with extensive testing and debugging of each intermediate architecture. Although Brooks does not claim that this incremental layering process is anything other than "a simplistic and crude analogy for evolutionary development" (Brooks, 1991a p. 1229), it is worthwhile considering whether there are similarities between this design process and the evolutionary path that has led to layered control systems in the vertebrate brain. In this section we therefore briefly outline some key findings and considerations concerning the evolution of the vertebrate nervous system focusing first on major morphological changes, and second on finer-grained changes in brain circuitry.

Major Evolutionary Changes in the Morphology of the Vertebrate Nervous System

Comparative and paleo-neurobiological studies reveal that a basic plan for the nervous system was established at a surprisingly early stage in vertebrate evolution (Jerison, 1973; Hodos, 1982; Miklos, 1993; Butler & Hodos, 1996). Specifically, the gross morphological divisions of the brain—*spinal cord, hindbrain, midbrain, diencephalon, telencephalon*—are present in all vertebrate classes and are also found in the earliest fossilized endocasts of jawless fish. This makes the general plan of the vertebrate brain at least four hundred million years old—indeed, it may have been in place as little as fifty million years after the Cambrian explosion which marked the first appearance of most of the modern metazoan phyla (Miklos, 1993). Many brain sub-divisions, including most of the components that make up the forebrain telencephalon, are also shared across the vertebrate classes and are therefore likely to have been present in early ancestors (Northcutt, 1981; Belekhova & Veselkin, 1985; Butler & Hodos, 1996). Figure 1 illustrates the major morphological divisions of the generalized vertebrate brain. All of the brain sub-divisions shown in the figure are found in all classes of vertebrates, with the exception of the *cerebellum*, which may be absent in jawless fish (Northcutt, 1996), and the *amygdala, septum*, and *striatum*, whose presence in jawless fish is suspected but not confirmed (Belekhova, 1990; Northcutt, 1994; Butler & Hodos, 1996). The forebrain *pallium* has three major sub-divisions in all vertebrates. In mammals, where the pallium is termed the *cerebral cortex*, most of the *dorsal* pallium forms the *neocortex*, while the *medial* pallium forms a group of structures that includes the *hippocampus*, and the *lateral* pallium forms the *olfactory cortex*. The extent to which each of these areas in mammals is homologous[4] with like-named pallial areas in non-mammals is, however, only partially resolved. There is insufficient space here for specific consideration of the various brain sub-divisions named in the figure, the reader is referred to Walker (1983), Arbib (1989), and Butler and Hodos (1996) for general introductions to brain function and evolution. A number of the brain regions shown are also discussed in later sections.

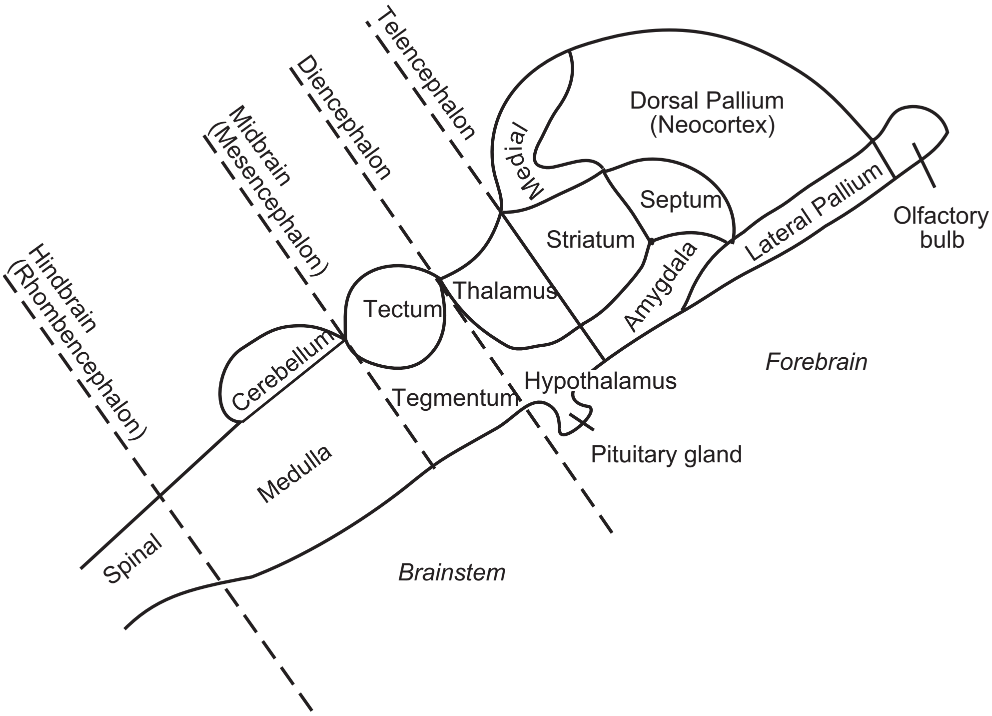

Figure 1. Diagrammatic representation of a generalized vertebrate brain showing the major morphological components (based on Nauta & Karten, 1970; Walker, 1983; Butler & Hodos, 1996).

Variations from the basic nervous system plan, in the different vertebrate classes, show specialization to different adaptive zones (ways of living) that in some cases have resulted in increased size and complexity of certain brain regions (Jerison, 1973; Harvey & Krebs, 1990). A well-known example is the massive expansion of the cerebral cortex and cerebellum in mammals[5] compared to ancestral synapsid reptiles, and the further expansion of these regions

---

[4] Homologous structures are those which have originated from the same structure in a common ancestor, since homology is defined by inheritance there may be derived differences in form and function.

[5] Less well-known, but of perhaps similar importance, is the expansion of various brain regions in vertebrate groups belonging to other classes. For instance, the telencephalon of birds is dramatically expanded compared to other non-mammalian tetrapods, as is that of some groups of sharks and rays relative to other cartilaginous fish, and, within the jawless fish, of hagfish relative to lampreys. These observations demonstrate the inadequacy of any

in some mammalian classes. Such changes do not, however, necessarily indicate the genesis of entirely new brain areas. For instance, the mammalian neocortex, perhaps the most derived structure in the mammalian forebrain, is likely to have evolved from areas in the dorsal pallium of earlier jawed vertebrates with similar incoming and outgoing connections (Northcutt, 1981; Northcutt & Kaas, 1995). Findings such as these have led researchers to conclude that vertebrate brain evolution has been a more conservative process than was previously thought. Expansion of forebrain centers in mammals (and other vertebrates with elaborated brains) has also not resulted in these regions taking over the functions of lower centers. Rather new or parallel functions have been added, complementing and modulating existing systems (Jerison, 1973; Sarnat & Netsky, 1981; Walker, 1983; Belekhova & Veselkin, 1985). A good illustration of this is provided by the evolution, in the mammalian nervous system, of direct connections from the neocortex to brainstem and spinal cord motorneurons. Such connections are not known in non-mammalian vertebrates. This development has been linked to the greater precision in digit use of many mammalian groups compared with the other amniotes (birds and reptiles) (see Butler & Hodos, 1996). It therefore seems likely that the new cortical motor pathways first evolved to provide precise and fine-grained control over manipulators, complimenting more rudimentary forms of motor control supplied by sub-cortical mechanisms. As noted earlier, a dissociation exists such that lesions in mammalian motor cortex leave these older, lower-level mechanisms for limb control intact.

Since a relatively stable brain architecture had already evolved in the first vertebrates, evidence for stages in the evolution of this basic plan must be sought in the elaboration of that system from the nervous systems of invertebrate ancestors. Vertebrates belong to the *chordate* phylum whose members all possess, at some stage in their development, a single, hollow nerve cord (the *neural tube*) which runs most of the length of the longitudinal body axis. Unfortunately, our knowledge of the evolution of the first vertebrate nervous systems from those of earlier chordates is largely speculative. There is no fossil evidence to speak of, and only one contemporary chordate, *Branchiostoma* (previously known as *Amphioxus*), that has a nervous system that resembles a possible transitional stage between other chordates and the vertebrates. Indirect evidence derived from the different branches of comparative biology is the main source of constraint for scenarios for the early evolution of the brain (e.g. Gans & Northcutt, 1983; Gans, 1989; Fritzsch, 1995; Butler & Hodos, 1996), while ongoing work tracing gene expression in neural tissue (e.g. Holland & Graham, 1995) represents the most promising prospect for resolving the many remaining arguments and uncertainties. Here, we briefly summarize a possible scenario (based on Gans & Northcutt, 1983; Holland & Graham, 1995; Northcutt, 1996) for the progression from an ancestral chordate nervous system to that of the first vertebrates.

The initial prevertebrate chordate was probably a filter-feeder with a segmented neural tube and an epidermal nerve net linking simple, unicellular sense organs. An intermediate stage involved the expansion of the neural tube to form the spinal cord and development of the hindbrain as a specialization of the most anterior segments of the tube. These changes accompanied the initial development of an anterior head with multicellular sense organs and probably coincided with a shift to a more active lifestyle based on predation. As the anterior sense organs became increasingly elaborated, integrative neural systems developed anterior to the hindbrain and formed the precursors of the vertebrate midbrain and diencephalon. The nervous system of *Branchiostoma*, though primitive in terms of the complexity of its sensory systems, contains possible homologues of various sub-regions of the hindbrain, and, more

---

view of brain evolution based on assumptions of a *scalae naturae* of increasing complexity from jawless fish to modern mammals (Hodos, 1982; Butler and Hodos, 1996).

controversially, of the midbrain and diencephalon (Lacalli, 1996; Northcutt, 1996), and may partially reflect this evolutionary stage. The telencephalon is generally believed to be a later addition to the prevertebrate brain although its evolutionary origins remain obscure.

The above observations allow us to consider some possible parallels between vertebrate brain evolution and the design process proposed by Brooks.

First, speculative scenarios for the early evolution of the vertebrate brain suggest that the more rostral (forward) brain centers may have been added to an ancestral nervous-system dominated first by the spinal cord (neural tube), and then by gradually elaborated hindbrain, midbrain, and forebrain components. This early phase of brain evolution was evidently more rapid than was previously thought and the basic architecture was probably in place with the appearance of the first jawless fish. Some aspects of this early brain evolution can perhaps be characterized as "incremental," for instance, it has been suggested that the hindbrain, which has clearly segmented structure, may have evolved by duplication and functional diversification of anterior segments of the ancestral neural tube (Metcalfe, Mendelson & Kimmel, 1986). In general, however, it has proved difficult, given the antiquity of these developments, to distinguish the addition of new brain structures from the co-option and adaptation of anterior structures already existing in an ancestral animal.

Second, the history of brain evolution within the vertebrate sub-phylum can be best understood as adaptation and divergence within the framework provided by the basic vertebrate brain plan. Some parallels with Brooks' view may exist, however, in the further evolution of the telencephalon in certain vertebrate groups, as exemplified by the expansion of cerebral cortex in mammals. Such changes may not be incremental in a strict sense (again, it is not clear to what extent they involve the genesis of new neural centers rather than the elaboration of existing ones), however, they demonstrate that where "new" higher-level motor and sensory systems have evolved these usually arise alongside existing centers which retain much of their original functionality.

Functional Change in Vertebrate Neural Circuitry

Over the course of evolution, successive radiations of vertebrates have invaded an increasingly diverse array of habitats and behavioral niches and, in so doing, have undergone several major transformations in morphology and behavior. In many new adaptive zones, the functional requirements that some neural circuits first evolved to satisfy will have altered or even disappeared, and the nervous system will have been required to support a variety of new perceptual, behavioral, and cognitive functions. A key issue which is still largely unresolved is the degree to which functional change has occurred by augmenting and modulating existing structures as opposed to radically reorganizing such structures. Here we summarize evidence that supports the first possibility—new functions arising without radical change to the existing nervous system organization.

A number of findings demonstrate that seemingly dramatic changes in function can take place through only minor alterations in cellular and connectional arrangements. For example, the spinal pattern generators that originally evolved to drive the undulatory movements of fish swimming have probably adapted to support the different stepping patterns observed in limbed vertebrates through a number of simple changes in the coupling and relative frequency of the network oscillators (Cohen, 1988). Wilczynski (1984) has suggested that major changes in peripheral sensory and motor apparatus may generally have been accommodated by a limited reorganization of central mechanisms. The flexibility to support such transformations arises, in part, through developmental processes that allow local areas to self-organize so as to accommodate change in circuits from which they receive inputs or to which they project outputs (see also Finlay, Wikler & Sengelaub, 1987; Deacon, 1990). There is also flexibility inherent in most neural circuits that allows them to be reconfigured into new modes of operation, without significant structural change, by modulation of their various cellular and synaptic constituents (Getting, 1989; Harris-Warrick, 1991). As a result, new functions may be obtained from existing circuits simply by the addition of appropriate modulatory inputs from other control centers. This mode of functional change has a striking parallel to Brooks' approach of obtaining new behaviors from established control circuitry by using subsumption mechanisms.

Kavanau (1990) has reviewed a number of arguments for the conservation of the redundant functions of neural circuitry through evolution. He claims that, as many neural circuits are multi-functional (see also Getting, 1989), natural selection will rarely operate to remove obsolete behaviors by physically excising the underlying neural pathways.[6] Instead, the elimination of behaviors is more likely to take place through the suppression or neutralization of redundant functions (using predominantly inhibitory connections), leaving the adaptive functions of the same neural circuits intact. The latent potential of brain circuitry to support redundant functions is shown in the observation of *relict* (i.e. ancient, no longer adaptive) behaviors that can be elicited in many vertebrates under circumstances of unusual stimulation, following brain damage, or during maturation (for examples see Kavanau, 1990). Again, in the suppression rather than excision of redundant functionality, we see a principle that brain evolution shares with the SA.

The above discussion has emphasized the conservative nature of neural evolution underlying functional change. It must be recognized, however, that there has been considerable change in the microstructural organization of many brain regions. This has included change in the size of cell groups, the addition of new cell types, and the migration of cell groups within the brain. Connectional changes have included increased differentiation of local areas through selective connection loss and axonal invasion of new areas. The relative importance and frequency of these different forms of microstructural change remains under dispute (Northcutt, 1984; Deacon, 1990; Fritzsch, 1995; Butler & Hodos, 1996), however, it is evident that significant reorganization of existing structures has taken place during vertebrate brain evolution. In this respect, Brooks' proposal to finalize the design of lower layers before higher layers are added represents a significantly more constrained way of building a control system than has taken place in the evolution of the vertebrate nervous system. Restricting change to the newest layer may prove a useful pragmatic constraint for first attempts at designing complex robot controllers. In the longer-term, however, the use of self-organization in the construction of control modules, and the development of methods that directly manipulate system dynamics (e.g. Schoner et al., 1995), might enable the design of behavior-based robots that retain more inherent flexibility at multiple levels of the evolving control system.

4 THE BEHAVIOR SYSTEMS APPROACH

One of the aspects of Brooks' SA that has had most impact is his proposal for a behavioral decomposition of the control system. Although there are earlier precedents for a behavior-oriented, and biologically-inspired, approach to robot control (for instance Walter, 1953), Brooks' work has been one of the main inspirations for the currently renewed interest in this

---

[6] This conservatism is a characteristic of central neural circuits, redundant peripheral systems, however, such as unused sensory circuits, are more likely to be uni-functional, and therefore are more likely to be eliminated in evolution (Kavanau, 1990).

field and has helped to initiate a more widespread move toward behavior-based artificial intelligence (Maes, 1992). The aim of this section is to briefly consider behavior-oriented control from a biological viewpoint. Specifically, we focus on work in ethology and psychology which has advocated a *behavior systems* approach to understanding animal behavior, and on neuroscientific evidence for behavioral decomposition in the architecture of the vertebrate nervous system.

The behavior systems approach in ethology is most closely associated with the work of Baerends, though it draws ideas and influences from many sources including the theoretical statements of Lorenz and Tinbergen (see references in Baerends, 1976). Baerends' approach begins with Lorenz's hierarchical hypothesis for the functional organization of behavior. This partitions the behavioral control of the animal into multiple levels with subordination of lower level "motivational states" by higher level ones, and with inhibition between states at the same level. Within this architecture, Baerends advocates a further decomposition into *behavior systems* which he defines as "relatively independent units" composed of "control systems of different orders [i.e. levels] fulfilling special tasks or sub-tasks in the context of a hierarchical organization" (Baerends, 1976, p. 736). At the systems level he has in mind decomposition into the major motivational systems such as defense, feeding, reproduction, body care, and so on. While at lower levels, each of these major systems breaks down into groups of related actions aimed at very specific behavioral goals. Figure 2 shows Baerends' example for the case of three interacting systems (and their sub-systems) in the behavior of the herring gull. Baerends emphasizes (as we have in Section 2) that his notion of hierarchy is sufficiently loose to allow feedback loops between different levels. He also recognizes interconnections and overlap between the various systems at multiple levels describing the whole as "an organization of interconnected networks" (Baerends, 1976, p. 731).

The behavior systems approach first emerged from ethological research and therefore provides a functional decomposition of animal behavior that need not map, in any direct fashion, onto the anatomical organization of the nervous system. Baerends expected, however, that the functional organization he described would somehow be reflected in the underlying neurophysiology. The literature on dissociations reviewed earlier suggests that the capacity to organize component movements into complex acts, and thence into motivated behaviors, involves increasingly rostral brain components. This indicates an interesting parallel between Baerend's hierarchy and the layered architecture of the vertebrate brain: higher-level brain circuits, by modulating and coordinating lower-level systems, appear to be necessary for generating the higher levels of functional integration recognized in vertebrate behavior.

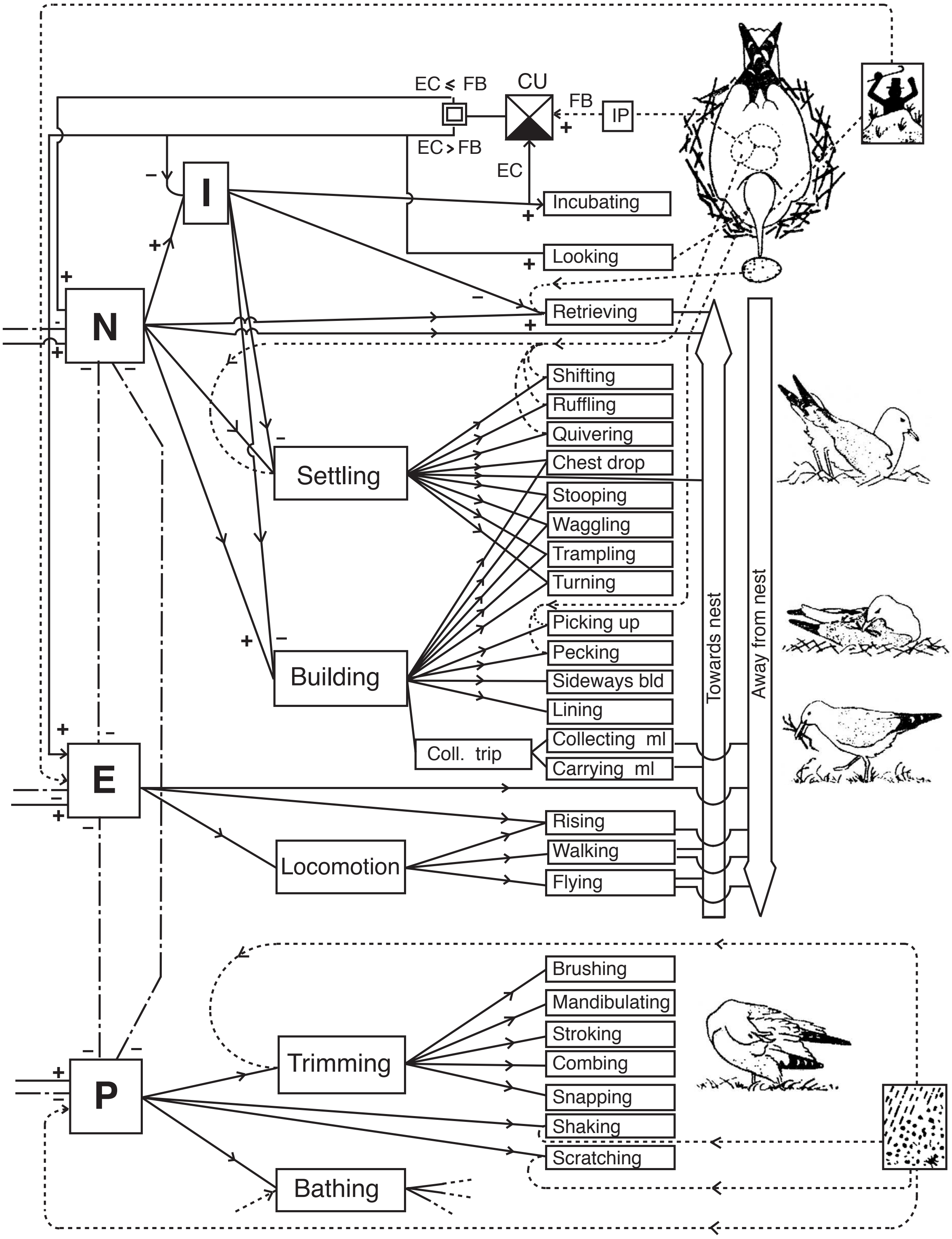

Figure 2. Model of behavior systems in the herring gull from Baerends (1970). The right-hand column shows the elementary behaviors or "fixed action patterns". To the left of these are the superimposed first and second order control systems (N= nesting system, E= escape system, P= preening system). The main behavior systems mutually inhibit one another. Reprinted from Baerends (1970) by permission of Brill Publishers.

There may also be physiological correlates for Baerends' decomposition of the control system into distinct behavioral units. Although, in general, the neural circuits underlying a specific behavior may be distributed and involved in implementing other functions, a degree of behavioral organization can be recognized in the vertebrate brain. For instance, the organization of parts of the spinal cord can be decomposed into simple movements (e.g. Bizzi, Giszter, Loeb et al., 1995), and parts of the midbrain and hindbrain into sets of coordinated movements (e.g. Depaulis, Keay & Bandler, 1992). A behavioral or movement-based organization has also been proposed for cortical motor systems, a suggestion that originated (once again) with Hughlings Jackson (1884/1958), although it remains controversial (see e.g. Humphrey, 1986; Wise, 1993). An example of behavioral decomposition in the nervous system will be given in Section 5 when we consider the organization of mid-brain defense systems in the rat.

An explicitly physiological interpretation of the behavior systems view has been adopted by Fanselow and Sigmundi (1987), who require that a behavior system, in addition to meeting certain ethological criteria, has an identifiable neural substrate. Further research relating the behavior systems approach to neurobehavioral and developmental studies has been reviewed by Timberlake (1993).

5 THE RAT DEFENSE SYSTEM VIEWED AS A SUBSUMPTION ARCHITECTURE

Brooks' SA provides a candidate architecture for distributed and layered robot control loosely inspired by evolutionary principles, while neuroscience shows empirical evidence for a layered control architecture in the vertebrate. Research in robotics, ethology, and neuroscience also suggests that behavior-oriented control may be a further valuable organizing principle for controlling a complex autonomous agent. To see if these parallels at a general level can be applied to a specific example of nervous system organization, this section looks in more detail at a specific vertebrate behavior system—the defense system of the rat. We focus on the rat defense system because it has undergone detailed scrutiny by researchers from a variety of disciplines, however, many of the general features we describe for this system are likely to have parallels in other behavior systems in the vertebrate brain.

The defense system of the rat can be viewed as a behavior system in the sense suggested by Baerends and others (Baerends, 1976; Fanselow & Sigmundi, 1987; Davey, 1989). Although widely distributed within the rat brain, and partly enmeshed with other neural mechanisms, it can therefore be considered as a relatively independent unit. Contemporary research in the neurobiology of rat behavior suggests that the defense system is organized as layered architecture in which higher level components provide increasingly sophisticated solutions to the problems of reducing and avoiding harm. The following summarizes this evidence and argues that the rat defense system shows some remarkable architectural similarities to Brooks' SA.

Layered Organization of Neural Mechanisms of Defense

The basic architecture of the rat defense system is illustrated in Figure 3. Each level in this architecture exploits a specialized subset of sensory processing mechanisms and connections to motor/autonomic outputs and, when activated by appropriate stimuli, is capable of delivering a set of appropriate adaptive reactions. The output of all levels appears to be relayed to a relatively restricted set of motor, autonomic, and hormonal effector units (Yardley & Hilton, 1986).

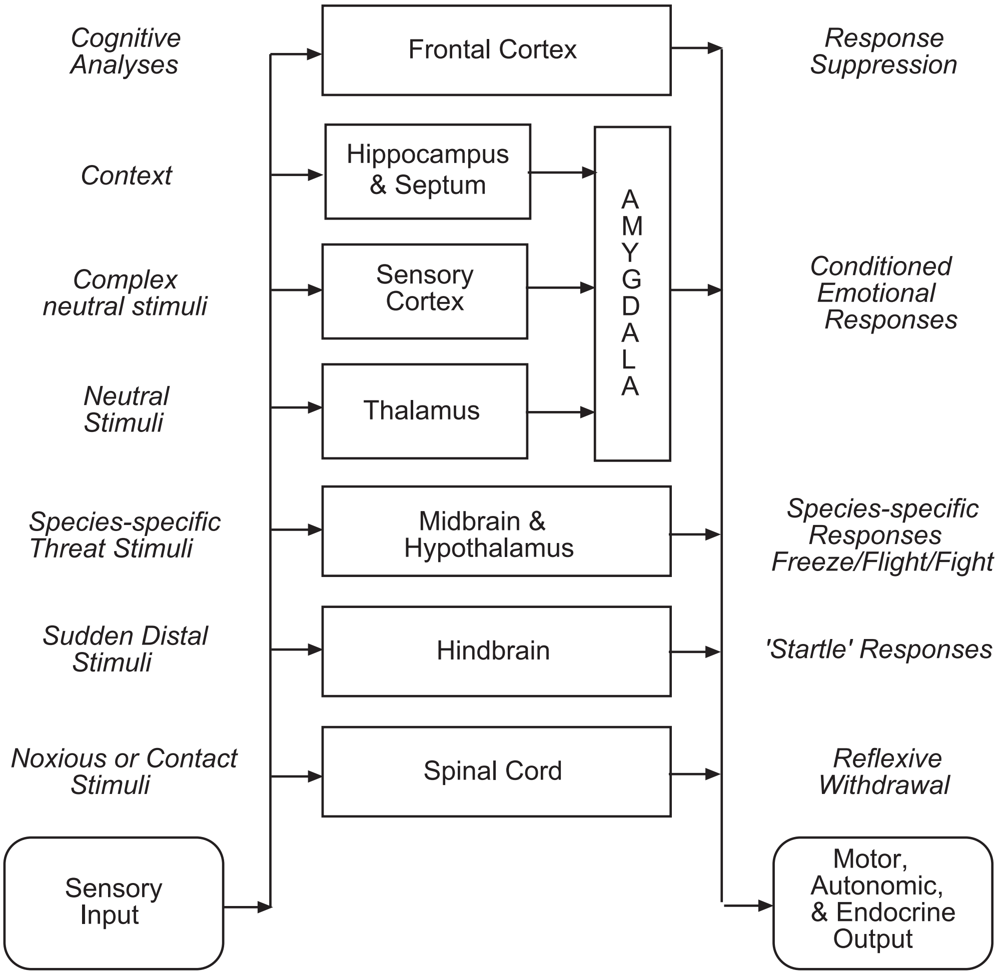

Figure 3. Layered organization of neural mechanisms of defense: higher level components provide increasingly sophisticated solutions to problems of reducing and avoiding harm. The nature of the sensory input, the principle brain structures involved, and the nature of the defense reaction are indicated for each level. See text for further explanation.

The following sub-sections describe the different layers of the rat defense system beginning with the lowest.

Spinal cord systems support avoidance of actual and imminent harm. The lowest level of the defense system deals with the problem of actual or imminent harm caused by noxious, contact stimuli. In the rat, as in other mammals, such mechanisms typically involve simple reflex withdrawal. For instance, if the rat's tail or paw touches something hot, a withdrawal reflex mediated at the level of the spinal cord removes the limb from the source of harm.

The capacity to perform similar escape responses probably existed in early chordates prior to the evolution of an anterior brain and multicellular sense organs. For instance, Gans (1989) has suggested that a simple “tail-flick” reflex, causing the sudden displacement of a floating animal, may have preceded the capacity for undulatory locomotion in the evolution of the prevertebrate motor system.

Hindbrain mechanisms provide rapid responses to stimuli that warn of possible harm. The sudden appearance of a visual or auditory stimulus is often linked with the possibility of physical harm. For a mammal, an adaptive response in these circumstances is to tense the somatic musculature. In the rat, a good example of this layer of the defense system is provided by the circuits that mediate the *acoustic startle response (ASR)* (Davis, 1984). These pathways, located entirely in the hindbrain and spinal cord, constitute the first interface between the anterior sense organs and the motor systems that deliver defensive behavior.

Hindbrain defense circuitry is highly conserved in vertebrates—for instance, the circuits underlying rat acoustic startle may be largely homologous to escape circuitry found in fish nervous systems (Lingenhohl & Friauf, 1994). Such mechanisms may have originated with the first truly mobile chordates as evolutionary developments that allowed sustained and co-ordinated forward locomotion encouraged the concentration of sensory systems and neural tissue at the anterior end of the animal (Gans, 1989; Fetcho, 1992).

Midbrain systems organize co-ordinated responses to harm. Intermediate level systems located primarily in the midbrain add further sophistication to both sensory and motor aspects of defensive behavior. Sensory systems provide for both the recognition and localization of distant, potentially threatening stimuli, while motor systems organize co-ordinated patterns of defense activity. This level provides the first means by which the animal can respond proactively to specific signs of danger.

A good example of an intermediate level defense mechanism is the acoustic sensory processing provided by the midbrain inferior colliculus which is considerably more discriminating than the hindbrain circuitry for the ASR. Sensory filters in the inferior colliculus[7] have evolved to recognize species-specific stimuli indicative of threat (Casseday & Covey, 1996), and depending on the characteristics of those stimuli, to trigger responses such as freezing and escape behavior (Brandao, Melo & Cardoso, 1993).

Vision can also provide early warning of potential danger. A large body of research demonstrates a key role for midbrain circuitry in visually-triggered defense behavior. For example, experiments reviewed by the Blanchards (1987) show that the distance to a sighted threat is critical in determining the defense strategy of a wild rat. Freeze is the characteristic response at long-range (> 10m), and flight at intermediate distances (5–10m) where an escape route is available, with speed of flight varying directly with the distance to the stimulus. When the distance to the threat is less than 0.5m the response changes abruptly from flight to defensive vocalization then, at even closer range, to a wild jumping and biting attack. Redgrave and Dean (1991) have suggested that a set of filters in the midbrain superior colliculus are differentially sensitive to the distance of a potential threat. According to this hypothesis, each filter is tuned to different levels and rates of darkening of the retina, and separately wired to motor systems delivering appropriate responses.

Compared with simple withdrawal or startle, escape and avoidance often involve the vigorous use of many muscles. This requires the development of: (a) systems that can co-ordinate the complicated movements involved in "flight" and "fight;" and (b) autonomic and endocrine systems that can sustain the increased energy requirements of vigorous activity. Motor circuits in the brainstem, for instance in the reticular formation and midbrain periaqueductal gray, provide for these co-ordinated patterns of muscular activity. The hypothalamus, although located in the forebrain diencephalon, is a further important

---

[7] The inferior and superior colliculi are the mammalian homologues of the midbrain tectum shown in figure 1.

intermediate level center involved in mid- to long-term hormonal and autonomic control of defense reactions. The capacity to continue a behavior in the absence of the original stimulus is vital for effective defense (an animal will not live long that stops running when the predator moves behind a bush!), and one of the functions of hormonal control initiated by the hypothalamus may be to maintain elements of the defense reaction after the disappearance of the initial trigger.[8]

Transection studies, in which all of the forebrain above the hypothalamus is removed (i.e. the limbic system and cortex) show that mammalian midbrain defense systems can be dissociated from higher level control systems (see reviews in Van Sommers, 1972; Gallistel, 1980). Released from the largely inhibitory control of forebrain systems these mechanisms will operate whenever the relevant triggers are presented regardless of whether they are appropriate in the wider context.

Homologues of mammalian midbrain systems will have been at least partially present in the first vertebrates (Northcutt, 1996) and are found with many ancient features still intact in all modern vertebrate classes (King & Comer, 1996; Butler & Hodos, 1996). Lacalli (1996) has argued that the nervous system of *Branchiostoma*, although tiny in size and cell number, contains homologues of the vertebrate eye, diencephalon, midbrain tectum, and brainstem motor areas. If correct, this would suggest that the intermediate level of organization described here may have originated in the nervous systems of the later chordate ancestors of the first vertebrates.[9]

Forebrain systems learn that arbitrary stimuli can be predictive of harm. At higher levels of the rat defense system, arbitrary events can be associated with harm and used as triggers for the full suite of defense reactions. The limbic system (the amygdala, hippocampus, and septum in Figure 1), and the amygdala in particular, appear to play a key role in the acquisition of these associations.

The amygdala receives sensory information of every modality alongside signals associated with pain and distress which arrive, probably directly, from the brainstem. Its outputs go to the various behavioral, autonomic and endocrine mechanisms through which defense reactions are commonly expressed. Classical conditioning in the amygdala allows a neutral stimulus (such as the presence of a light) to be associated with pain signals (such as a harmful footshock), such that after a few pairings, defensive reactions are evoked by the neutral stimulus alone (see Ledoux, 1995). This is termed a *conditioned emotional response*. The amygdala therefore appears to operate by putting an “emotional stamp” on any sensory input which is present at the same time, or occurs just before, something nasty. When the same stimulus occurs again it is given immediate access to the defense circuitry. By linking aversive triggers with formerly neutral cues, this mechanism allows a defense reaction to be continued until the animal has moved to a safer environment. The amygdala therefore provides a more refined means of controlling long-term defense behavior than is possible through midbrain and hypothalamic mechanisms alone.

For simple stimulus features, the ability to acquire a conditioned response appears to be dissociated from the highest cortical centers. For example, subcortical sensory structures can

---

[8] In situated robotics the device of “time-outs” is often used for a similar, if cruder, effect.

[9] Indeed, taking *Branchiostoma* as a model may underestimate the complexity of later prevertebrates. Being a filter-feeder rather than a predator it lacks the well developed sense organs and associated neural systems that such creatures are likely to have possessed (Gans, 1989).

provide, via the thalamus, the input for the conditioned emotional response to a simple tone or light (see Davis 1992; Ledoux, 1995). Learning, in both these circumstances, is unaffected by lesions to auditory or visual cortex.

Cortical systems provide additional tiers of input to the amygdala. Further analyses of significant environmental events are performed in sensory cortex and in cortical association areas—structures which also have direct output projections to the amygdala. The hippocampus and septum also project to the amygdala providing access to additional contextual knowledge that can help the animal to distinguish between threatening and non-threatening situations. Ledoux (1992) suggests that thalamic, cortical, and hippocampal areas provide the amygdala with "three tiers" of sensory and contextual information.

The limbic system (including homologues of the mammalian amygdala) is present in all vertebrates although data on the earliest vertebrate class (the jawless fish) is sparse (Belekhova, 1990). It therefore seems likely that the role of the amygdala in assessing the biological significance of sense data, and modulating lower level behavioral and autonomic sub-systems, was established early in vertebrate evolution. The amygdaloid nuclei become differentiated in the transition from amphibians to reptiles, and undergo further, though less substantive, change in the transition to mammals (Bruce & Neary, 1995). These developments appear to give the expanded forebrain cognitive and sensory sub-systems of later vertebrates access to the same basic machinery for acquired emotional responses.

Frontal cortex may unlearn associations between arbitrary stimuli and harm. Such are the vagaries of life that an acquired association between a stimulus and harm can later cease to be appropriate. LeDoux (1995) has suggested that a yet higher level of competence is required to *unlearn* some associations between arbitrary stimuli and harm. This proposal arose from experiments on the conditioned emotional response that began by pairing a tone with an electric shock. After relatively few trials, defense reactions to the tone were well established, at which point the shock was turned off so that the tone no longer predicted the aversive stimulus. Very gradually, over several weeks, the conditioned emotional reactions to the tone subsided and eventually disappeared. Once extinction had occurred, restricted lesions were placed in the frontal cortex. On the first trial following these lesions the conditioned emotional response (i.e. to the tone alone) re-emerged at full strength.

This remarkable result suggests that acquired associations between conditioned and unconditioned stimuli at the level of the amygdala remained intact throughout the extinction period. Instead, it appears that the frontal cortex gradually learned that the tone was no longer dangerous and established an appropriate inhibitory connection to the amygdala. When the frontal cortex was lesioned, this modulating inhibitory input was lost releasing the emotional link between the tone and defense reactions. This form of extinction thus appears to involve not the eradication of a learned response but *active inhibition of that response by a higher level system*.

This component of the defense system has been demonstrated only in the mammalian forebrain. However, other highly encephalised vertebrate groups (e.g. birds) may have evolved similar mechanisms.

Parallels between the Rat Defense System and the Subsumption Architecture

The above review paints a fascinating picture of the defense system in the rat brain. The similarities to a subsumption architecture are clear and can be summarized as follows:

Distributed control. The defense system is composed of multiple, vertically organized, sub-systems operating in parallel. Each level makes an important contribution to the functionality of the whole system. If the higher layers of the system were to “take over” to the extent that lower level mechanisms became more or less vestigial, then this would be permissible under an SA, but is clearly against the spirit of Brooks’ proposal. This does not seem to be true of the defense system as described above—the different layers make distinct and important contributions.

Behavioral decomposition. Although much of the neural circuitry in the vertebrate brain appears to be multifunctional, there is some evidence of dissociations between the defense system and other behavior systems in the rat brain, and of a finer-grained behavior-based organization within the defense system itself. Both of these points are illustrated by research on visual response circuitry in the rat midbrain.

First, studies of the superior colliculus reviewed by Dean, Redgrave and Westby (1989) indicate the presence of at least two separate visual response systems in rat midbrain, one specialized for orienting movements such as tracking and pursuit, the other for defensive behaviors such as avoidance and flight. These two systems appear to rely on separate and dissociable pathways to brainstem/spinal cord motor systems. Similar findings in frogs (King & Comer, 1996) suggest that this separation may be a general characteristic of the vertebrate midbrain.

Second, within the midbrain defense system, evidence for a behavior-based organization of defense reactions is found in the lateral midbrain periaqueductal gray—the intermediate level brain structure responsible for coordinating many lower-level defense reactions. For instance, backward defense behavior, characterized by backing off and defensive rearing, and forward defense behavior, characterized by forward locomotion and occasional jumping, appear to be supported by two neighboring but distinct populations of neurons (Depaulis et al., 1992). As we have already suggested, these two contrasting defense responses may be triggered by separate filters in the superior colliculus tuned to different forms of visual threat (Redgrave & Dean, 1991).

Increasing levels of competences. At ascending levels the defense system show increased ability to anticipate danger and greater complexity of response. Lower level defense reactions do not depend on higher level structures and will operate when triggered unless actively inhibited by higher levels. There is a dissociation between levels such that damage to higher level structures leaves lower level behaviors relatively unimpaired.

Incremental construction. As we pointed out in Section 3 the antiquity of the basic vertebrate brain plan makes it difficult to establish definitively whether the finding of a layered nervous system architecture is indicative of a phylogenetic sequence. We have noted above that all but the highest levels of the defense architecture are shared across all the vertebrate classes but may have evolved in a loosely incremental fashion in prevertebrates. There is generally good evidence of conservation of lower level defense circuitry through evolution.

Communication between levels by subsumption mechanisms. Anatomical studies indicate that outputs from higher level modules such as the amygdala and cerebral cortex converge on lower somatomotor, autonomic and endocrine output units. Thus higher level systems act by modulating outputs generated at lower layers. In the rat defense system we see several examples of subsumption-like mechanisms that implement this modulation.

First, higher-level systems can *coordinate* multiple lower level mechanisms. This is observed in the coordinated control, by mid-brain defense sub-systems, of sets of lower-level reflex and pattern-generating mechanisms.

Second, higher-level systems can provide *substitute* inputs to lower-level output mechanisms in the absence of primary noxious sensory information impinging on the lower layer. This is evident in the conditioned emotional response where low-level defense reactions can be triggered by arbitrary and potentially complex configurations of non-noxious sensory stimuli. A further example is the task of orienting to the position of a remembered, and previously dangerous, target. In such situations, high level systems clearly must impose control without low-level sensory input. Thus, the low-level gaze positioning system (the other principle function of the superior colliculus) operates regardless of whether its input denotes a direct stimulus or a remembered object (Wurtz & Hikosaka, 1986).

Third, higher level systems can *inhibit* those below. This is essential if one is to avoid assaulting the dentist while he/she is drilling your teeth! A more formal example, described above, is the continuous inhibitory signal[10] from the frontal cortex to the amygdala that allows unlearning of the conditioned emotional response.

Finally, lower level reactions can be modulated by simple mechanisms such as *gain* control.[11] An example of this is the *potentiated acoustic startle response* (Davis, 1992) in which the amplitude of a low level reflex behavior is dramatically enhanced when the animal is in a “fearful” state initiated by a higher level module. This *potentiation* of the startle response is caused by an output from the amygdala that taps into the lower level reflex arc.

Limited sensor fusion at lower levels, no central models. Signals from the different sensory modalities converge on motor and pre-motor defense mechanisms throughout the brain. This capacity to respond to threat stimuli from multiple modalities does not imply, however, that signals from the different senses are always combined to form unified representations. Indeed, in the context of brainstem defense sub-systems it seems more likely that the modalities supply relatively independent triggers or combine in a fashion that depends on fairly simple stimulus properties such as intensity. Where there is evidence for more sophisticated forms of sensor fusion, higher levels of the nervous system are usually implicated. For instance, in mammalian brains, near–threshold visual and auditory threat signals from sources that correspond in place will combine to form a more effective trigger than either stimulus presented alone or both stimuli presented in different places (Meredith & Stein, 1983). In other words, seeing the branch twitch and hearing the leaves rustle will increase the likelihood of a reaction, but only if these two events occur *at the same location*. Wallace and Stein (1997) have recently shown that this solution to the spatial correspondence problem critically relies on circuitry in the associative cortex—animals in which the relevant cortical areas are temporally inactivated, and infant animals in which cortex is not fully matured, fail to show this cross-modal sensory enhancement.

The dissociations, noted above, between the different layers of the defense system demonstrate that separate sensor processing channels exist at multiple levels of the brain. Lower-level channels, which may exploit relatively fast, task-specific, pattern recognition mechanisms (such as the sensory filters in the superior colliculus) appear to be independent of representations established in higher layers, Brooks’ notion of “abstraction barriers” is thus at least partially respected.

---

[10] It is interesting to note that this exact form of subsumption control (continuous, active inhibition) was a refinement that was added to the original SA (Brooks, 1989) to replace inhibitory mechanisms based on time-outs.

[11] This can be implemented in the subsumption architecture by a high-level input that subsumes the role of a constant gain parameter in the lower-level system.

In summary, the above correspondences would appear sufficient to justify our view that the architecture of the defense system shares many interesting characteristics with Brooks' SA.

6 ACTION SELECTION IN LAYERED CONTROL ARCHITECTURES

A problem of conflict resolution arises whenever multiple functional units are in competition for a limited set of effectors. At any time only one directing system can generally be allowed control of a given set of motor outputs. In the adaptive behavior literature this situation is commonly referred to as an *action selection* problem (Tyrrell, 1992; Maes, 1995). A frequent criticism of Brooks' SA by authors otherwise sympathetic to the behavior-based approach concerns the inflexible nature of the conflict resolution that can be achieved by subsumption mechanisms (Rosenblatt & Payton, 1989; Maes, 1990; Blumberg, 1994; Schoner et al., 1995). Brooks has noted that the problem of designing an appropriate subsumption scheme becomes harder as more control layers are added and has investigated a number of extended architectures that include mechanisms designed to improve action selection (Brooks, 1995). This is clearly an area where insights from natural control systems could prove useful in robot design.

The problem of resolving conflicts between competing behaviors has been an important focus of research in ethology (see. e.g. McFarland, 1971; Colgan, 1989), which has contributed a number of important conceptual and control-theoretic models, some of which have been evaluated in computer simulations (e.g. Ludlow, 1976; Houston & Sumida, 1985). Recent research in adaptive behavior has partly drawn on these earlier ethological models and has also moved towards evaluating candidate action selection mechanisms in more complete simulations of the agent and its environment (e.g. Tyrrell, 1992; Blumberg, 1994; Maes, 1995) or in the control architecture of a genuine mobile robot (Snaith & Holland, 1990; Brooks, 1994; Rosenblatt, 1997; McFarland & Spier, 1997). Only limited attention has been paid, however, either in the ethology or adaptive behavior literature, to the neural circuitry that supports conflict resolution in vertebrate nervous systems. An understanding of the functional architecture of such circuits should provide a useful additional source of constraint for designing effective, biologically-inspired action selection mechanisms.

Throughout most of this article we have argued that the vertebrate brain employs vertical, subsumption-like mechanisms to implement a layered control architecture. In Section 2, however, we noted that the *basal ganglia* (BG), a group of functionally-related, central brain structures, appear to have a controlling influence on neural systems at multiple levels and so form an important exception to the overall vertical decomposition of the brain. We will argue below that the core role of the BG is as a specialized action selection mechanism that provides a more flexible means of conflict resolution than is possible through subsumption mechanisms alone. Before outlining this proposal, however, we briefly review some of the requirements for effective action selection and consider whether there is a need for specialized selection circuitry within a complex control architecture such as the brain.

The Action Selection Problem

The problem of action selection has been considered at length elsewhere (McFarland, 1989; Snaith & Holland, 1990; Tyrrell, 1992; Maes, 1995). The following therefore focuses on a number of issues of particular relevance to the subsequent review of vertebrate conflict resolution mechanisms.

A basic principle of action selection is that from a set of *incompatible* competitors only one should be allowed expression at a given time. Further requirements for effective action selection can be loosely divided into those related to making appropriate selections and those concerned with effective switching between competitors.

With respect to selection, a widely applied heuristic—and one that appears to be exploited in vertebrate decision-making (McFarland, 1989)—is to prefer the most strongly supported competitor as indicated by relevant external and internal cues. The capacity for a specific competitor to prevail in a selection contest should be weighted to provide appropriate dominance relationships between different competitors. So, for instance, safety-critical competitors (such as defense behaviors) should be able to dominate even in the presence of weak triggering stimuli. In natural control systems we would expect this balance between competitors to be adaptable through learning as well as through natural selection.

In order to facilitate effective and timely *switching* between competitors, we can identify (following Snaith & Holland, 1990) a number of useful properties that an action selection mechanism should possess. First, a competitor with a slight edge over its rivals should see the competition resolved rapidly and decisively in its favor so providing *clean switching*. Second, the presence of other competitors which are activated but not engaged should not interfere with the performance of the winning behavior once the competition has been resolved. This can be termed *lack of distortion*. Both these properties can be provided by mechanisms of a winner-takes-all variety. We therefore may expect to find neural circuits that implement winner-takes-all involved in action selection in the vertebrate brain. Finally, it may also be useful for a winning competitor to remain active at lower input levels than are initially required for it to overcome the competition. This characteristic, termed *hysteresis* (Snaith & Holland, 1990) or *persistence* (McFarland, 1989), can prevent unnecessary switching and is often implemented through some form of non-linear, positive feedback loop (see e.g. Houston & Sumida, 1985).

Conflict resolution between competitors bidding for incompatible uses of a single resource is only part of the wider problem of generating coherent behavior (Brooks, 1994). Different effector systems, such as the muscle groups underlying locomotion and gaze in mammals, constitute more-or-less independent resources, however, it is clearly important that their activities are appropriately co-ordinated. For instance, it is generally considered unwise to persist in walking forwards while looking backwards! As this example demonstrates, selection mechanisms for individual resources need to be embedded within a control architecture that can deliver appropriate simultaneous and sequential patterns of activity in multiple output systems.

Is There a Need for Specialized Selection Circuitry in Complex Control Architectures?

Emergent action selection. Work in the field of adaptive behavior has been at the forefront of the study of *emergent functionality* (see, e.g. Steels, 1995) where useful behavioral outcomes are seen to arise as a consequence or side-effect of the interaction of control system components which individually have a different or more limited functionality. In such systems no one component is decisive in shaping the overall outcome, and it is generally impossible to attribute specific aspects of behavior to the function of individual components. Various authors have proposed that effective action selection can arise as an emergent function of a suitably configured control system (for review see Maes, 1995). For instance, Maes (1990, 1995) has described a control architecture for an "artificial creature" which is decomposed into multiple goal-oriented behaviors. In this system there is no specific locus at which action selection takes place, rather a network of excitatory and inhibitory connections between behaviors is configured so as to allow an appropriate sequence of actions to be generated as the agent interacts with its world. Mutual inhibition between incompatible behaviors ensures that they are not allowed simultaneous expression. Maes argues that such architectures can allow flexible action selection, of a natural and fluid character, that is robust to damage to individual components of the system. As discussed by Kelso (1995), emergent selection is also demonstrated in multi-functional pattern-generating networks (such as those underlying vertebrate locomotory behavior) that can be driven by small changes in input (or in neuromodulatory substances) to switch from one behavioral mode to another. The intrinsic dynamics of these circuits allow clean switching between motor patterns without the need for identifiable switching elements. The methodology of dynamic systems modeling is proving important in understanding how appropriate, emergent switching behavior can be derived from such multi-component control circuits (Kelso, 1995; Schoner 1995).

Specialized action selection. The alternative to emergent selection is to allocate specific components and pathways in the control system to the task of resolving selection conflicts. A general argument in favor of such specialized circuitry concerns the advantages of modularity in the design of control systems. To the extent that the problem of selection can be distinguished from the perceptual and motor control problems involved in coordinating a given activity it should be advantageous to decouple the selection mechanism from other parts of the control circuitry. As separate components each can be improved or modified independently. In contrast, in a circuit that displays emergent selection, a change directed at some other aspect of function could impact on the switching behavior of the network with possibly undesirable consequences. The advantages incurred by modularity in dissociating functionally distinct components of the system are probably as significant for evolved systems as they are for engineered ones (Wagner & Altenberg, 1996).

Specialized action selection mechanisms feature in the behavior-based robot control system described by Rosenblatt (1997), which incorporates a central arbitration mechanism in an otherwise distributed architecture. Below we summarize evidence indicating that for some aspects of selection the vertebrate brain has evolved specialized switching mechanisms that are largely decoupled from other aspects of control.

Distributed selection mechanisms—recurrent reciprocal inhibition. A specific form of neural connectivity, which is often associated with emergent selection, is mutual or reciprocal inhibition. At its simplest, reciprocal inhibition is the process, identified by Sherrington (1947), whereby a signal that excites a muscle on one side of a joint inhibits the antagonistic muscle on the opposite side. A more complex form arises in networks with *recurrent reciprocal inhibition* (RRI) (see e.g. Gallistel, 1980) where two or more units are connected such that each one has an inhibitory link to every other (see Figure 4, left). Such circuits display a form of positive feedback since increasing the activation of one unit causes increased inhibition on the remaining units thereby reducing their inhibitory effect on the first. RRI can therefore support winner-takes-all switching making it an attractive means for implementing action selection (Snaith & Holland, 1990). The strengths of incoming excitatory links and of the inhibitory links between units can also be tuned to support a complex pattern of dominance relationships between competitors.

RRI is one of principles exploited in Maes (1990) action selection architecture for an artificial creature, it has also featured in a number of ethological models of action selection in invertebrates (e.g. Ludlow, 1976; Edwards, 1991) and vertebrates (see Tyrrell, 1993). RRI connectivity has been identified in many different areas of the brain (Windhorst, 1996), however, its functional role in most situations is not clearly understood. Gallistel (1980) has suggested that such circuits could play a role in conflict resolution at multiple levels of the vertebrate nervous system.

The distributed nature of RRI means that selection can be viewed as an emergent property of the network (Snaith & Holland, 1990). Whether such circuits display emergent selection in a stronger sense will depend, however, on whether the reciprocal inhibitory connections that implement switching have any further influence ongoing behavior. If it can be shown that selection components of such a circuit are dissociable from those involved in generating specific behavioral outcomes then the RRI circuitry could be described as implementing a specialized, albeit distributed, selection mechanism.

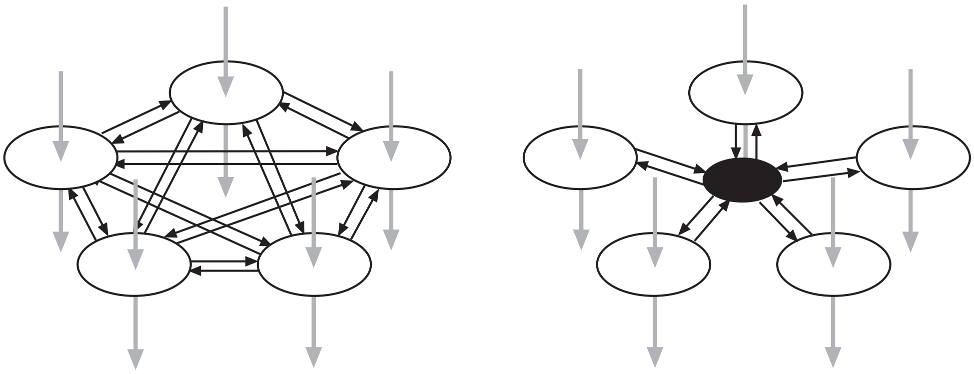

Figure 4. Distributed recurrent reciprocal inhibition network (left), compared with a "central switch" selection mechanism (right). The latter has a significant advantage in terms of economy of connection costs. Dark arrows indicate one-way inhibitory connections, light arrows indicate sensory inputs and motor outputs.

Centralized action selection mechanisms. Snaith and Holland (1990), following McFarland (1965), have contrasted RRI action selection with a system that employs a specialized, central switching device (see Figure 4, right). They note that to arbitrate between *n* competitors, an RRI system with full connectivity requires *n(n-1)* inhibitory connections, while adding a new competitor requires a further *2n* connections. In contrast, a system which uses a central switching device to host and resolve the competition requires only two connections per competitor (to and from the switching mechanism) resulting in *2n* connections in all. Adding a further unit requires only *two* additional connections. Snaith and Holland argue that mechanisms of both kinds can, in principle, exhibit clean switching, minimal distortion, and hysteresis. Clearly, however, a central switching device provides a significant advantage in terms of economy of connections costs. Except for the overhead of the switching device itself, such a system is cheaper to build and maintain than an RRI system. Perhaps as importantly, it is also cheaper and easier to increment, so simplifying the task of integrating new competitors into an existing system.

Ringo (1991) has pointed out that geometrical factors place important limits on the degree of network interconnectivity within the brain. In particular, larger brains cannot support the same degree of connectivity as smaller ones—significant increases in brain size must inevitably be accompanied by decreased connectivity between non-neighboring brain areas. Leise (1990) has argued that a common feature of both vertebrate and invertebrate nervous systems is that they are composed of anatomically and functionally differentiable local compartments which are restricted in size to a maximum of around 1mm diameter. Connectivity between neurons is highest within compartments, and larger nervous systems have more compartments rather than larger individual compartments. One of the constraints that appears to limit compartment size is the greater cost of high-bandwidth communication over long distances in neural tissue.

The nature of the action selection problem is such that functional systems in different parts of the brain will often be in competition for the same motor resources. The requirements of lower connectivity and increased compartmentalization with increased brain size therefore strongly favors selection architectures with lower connectional overheads. In the remainder of this section we propose that the vertebrate brain exploits specialized action selection mechanisms that are instantiated within the basal ganglia. The centralized nature of these mechanisms provides an efficient means of arbitrating between functional units that are widely distributed in the brain. We will suggest, however, that within the BG distributed switching mechanisms are employed in a manner that exploits their capacity for effective switching whilst minimizing connectivity costs.

Structure and Function of the Vertebrate Basal Ganglia viewed as an Action Selection Device

The principle components of the vertebrate BG are the *striatum* and *pallidum* found at the base of the forebrain and the *substantia nigra* and *ventral tegmental area* in the midbrain. These structures appear to have homologues in the nervous systems of all classes of jawed vertebrates (Medina & Reiner, 1995; Butler & Hodos, 1996), and possibly in all vertebrates (Northcutt, 1994). Figure 5 shows the main BG nuclei and some of their intrinsic and extrinsic connections within the mammalian brain.

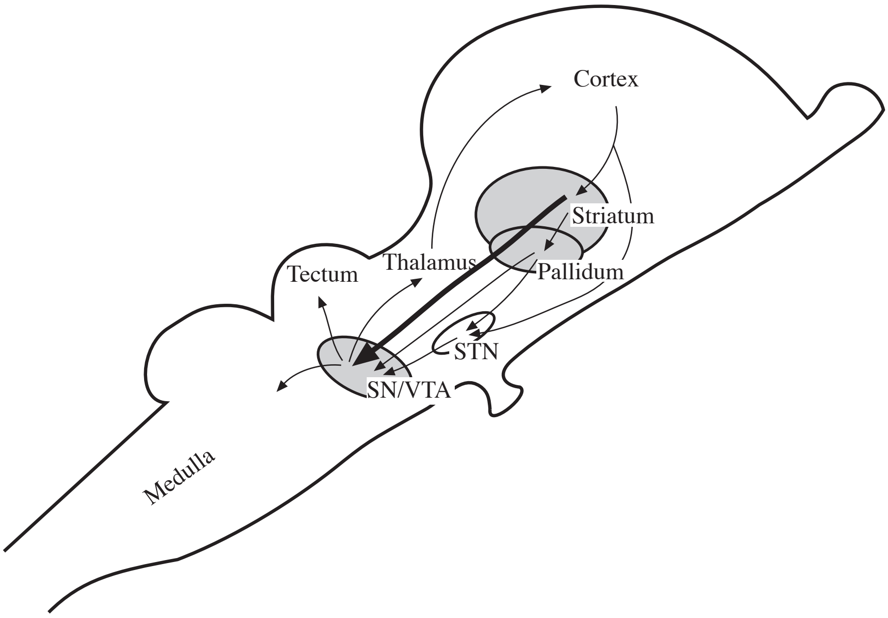

Figure 5. Principle structures of the basal ganglia and some of their intrinsic and extrinsic connections in the mammalian brain. Based on Medina and Reiner (1995). Abbreviations: SN/VTA—substantia nigra/ventral tegmental area, STN—subthalamic nucleus. The direct striatonigral pathway (shown by the large arrow) is found in all jawed vertebrates, and possibly in all vertebrates.

The proposal that the BG performs action selection in the vertebrate brain is not a radical perspective on BG function but rather derives from a growing consensus that a key function of these structures is to enable desired actions and to inhibit undesired, potentially competing, actions (for reviews see Chevalier & Deniau, 1990; Salamone, 1991; Reiner, 1994; Groenewegen, Wright & Beijer, 1996; Mink, 1996; Redgrave, Prescott & Gurney, in press). This literature suggests the following view of the functional architecture of the BG.

Activity relating to "bids" for access to motor apparatus appears to be continuously projected to the input side of the BG from relevant functional sub-systems in both the brainstem and forebrain of the animal. This activity may form the "common currency" in which competing requests for actuating systems can be effectively compared. Internal circuitry within the BG then determines a "winner" whose contact with the motor systems is specifically disinhibited via its output channel.

The following briefly summarizes some of key anatomical, neurophysiological, and neurobehavioral findings in support of this proposal.

Inputs to the BG include branched pathways from sensorimotor systems and contextual inputs from sensory and motivational areas. Anatomical evidence shows that cortical and brainstem sensorimotor systems, plus several of the forebrain limbic structures, communicate directly with motor and pre-motor mechanisms in the brainstem and spinal cord. However, these systems also project, usually via a collateral (split) pathway, to the *striatum*, the main input center of the BG. This branch could allow them to enter into a competition for control of the motor outputs hosted within the BG. Afferents from a wide range of sensory and motivational systems also arrive at BG input neurons. These connections could allow both extrinsic and intrinsic factors to enter into a "vast machinery" of context-specific filters in the striatum (Mink, 1996), influencing the strength of rival bids, and hence the currently preferred course of action. The input connectivity of the BG therefore indicates that it is well placed to resolve the problem of selecting an appropriate action for a given set of circumstances.

Outputs of the BG selectively inhibit and disinhibit movement generators. A tonic (i.e. continuous) inhibitory signal emerges from the main output centers of the BG (parts of the substantia nigra, ventral tegmental area, and pallidum), directed at neural circuits that either directly or indirectly generate movement. This inhibitory signal places a powerful brake on these movement systems. Commands emanating from the striatum inhibit the inhibitory BG output centers thereby *disinhibiting* selected movement generators. In the absence of such commands there can be no voluntary movement. The BG thus seems to hold a "veto" over midbrain and forebrain systems that seek access to the motor system; this is relinquished, for a selected action, through the mechanism of disinhibition.

Projection lines through the various sub-components of the basal ganglia appear to be largely organized into segregated parallel "channels." This segregation is maintained in the disinhibitory projections to output targets. Behavioral studies indicate that although the architecture of these channels is similar throughout most of the BG, different areas are functionally heterogeneous. For instance, there appears to be a motor somatotopy in parts of rat striatum such that restricted lesions at different locations effect different actions such as forelimb manipulation, biting, and gait (see, e.g. Cho & West, 1997). This would suggest that the circuitry in these local areas of the BG may primarily be used to resolve conflicts between systems bidding for incompatible uses of specific muscle groups. More generally, each local group of parallel circuits in the BG may be competing for a single output mechanism thereby forming a single, multi-way “switch.” If this interpretation is correct then the BG as a whole may provide an array of similar switching devices.

Winner-takes-all mechanisms in the BG could support effective switching. The above describes the principle “selection” pathway through the BG. However, the various BG nuclei have a rich interconnectivity, partly illustrated in Figure 5, whose function is not well understood. For instance, some intrinsic BG pathways have a net-excitatory effect on BG output nuclei that results in increased inhibition (rather than disinhibition) of motor mechanisms. Mink (1996) has reviewed anatomical and electrophysiological evidence from the mammalian BG indicating that the different intrinsic pathways provide complimentary mechanisms that act to focus activity in the BG output nuclei. This evidence suggests that the BG could implement a winner-takes-all competition between multiple channels—disinhibiting a desired channel while maintaining or increasing inhibition on competing channels. This proposal that has been explored and developed in our own recent computational model of BG intrinsic circuitry (Gurney, Prescott & Redgrave, 1998). If this interpretation is correct, it would appear that the BG is appropriately “wired” to provide clean switching between actions and lack of distortion in a selected act. A summary of this model of the BG as an action mechanism is shown in Figure 6.

It is interesting to note that RRI circuitry within the striatum probably provides one of the mechanisms that resolves the selection competition (Wickens, 1997). The axon collaterals of individual striatal neurons typically occupy an area of about 0.5 mm in diameter, so local RRI competitions can be expected to occur in regions of approximately this size. Note that this is below the 1mm limit suggested by Leise (1990) for compartment size (and therefore low-cost, high bandwidth communication) in nervous systems. By implementing RRI within confined and very local areas of the striatum the vertebrate brain may have evolved to exploit the potential of RRI switching circuits whilst minimizing the cost of inter-competitor connectivity—employing a micro-scale distributed switch within a specialized and centralized selection mechanism.

Behavioral effects of BG interventions and damage support a role in behavior selection and switching. Lesions, and neurochemical or electrophysiological interventions in BG structures have been shown to effect various aspects of behavior selection and switching in a number of different experimental paradigms (see e.g. Gelissen & Cools, 1988; Bakshi & Kelley, 1991; Salamone, 1991; Salamone et al., 1997). Depending on the site and nature of the intervention, these effects include changes in the dominance relations between behaviors, reductions and increases in switching relative to controls, changes in variability of behavior, and failure to complete behaviors. The BG therefore appear to be heavily implicated in maintaining appropriate frequency and timing of behavior switches in the normal state.

The symptoms of various brain disorders associated with the BG also suggest a role in action selection. Human disorders in which the BG is implicated include: Parkinson’s disease, whose symptoms include slowness and difficulty in making voluntary movements, Huntington’s disease and Tourette’s syndrome, characterized by excessive or inappropriate movement, and obsessive-compulsive disorders marked by repetition of certain stereotyped activities. Recent evidence suggests that these disorders may be related to abnormal levels of the neuromodulator *dopamine* in the BG. An important role of striatal dopamine may be to moderate the balance between the different control pathways through the BG (see e.g. Reiner, 1994). A deficit of dopamine, such as is seen in Parkinson’s, appears to favor the net-inhibitory pathways within the BG leading to too much inhibition on motor systems. Excess dopamine, on the other hand, may favor the disinhibitory (net-excitatory) pathway leading to the excessive movement seen in Huntington's disease and Tourette's syndrome, or causing certain activities to become over-dominant as seen in obsessive-compulsive disorders.

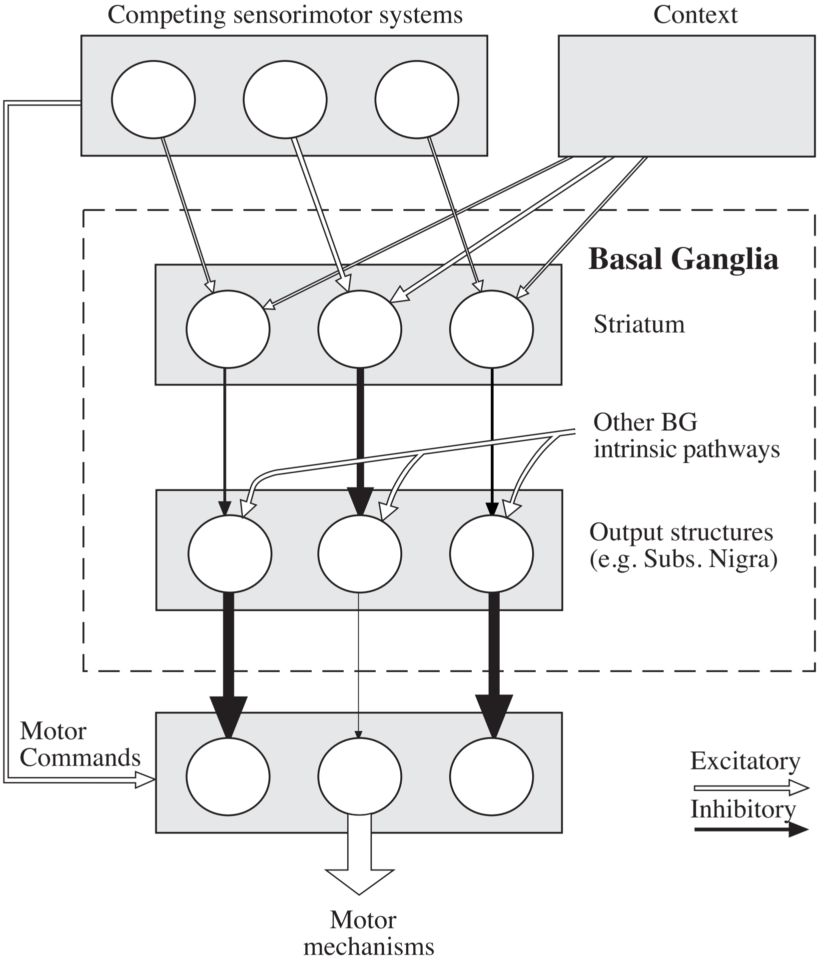

Figure 6. The BG viewed as an action selection mechanism. The competition between multiple sensorimotor mechanisms is resolved in the BG. The winning competitor inhibits tonically active neurons in BG output structures which disinhibits required motor mechanisms. Other BG intrinsic pathways excite output neurons contributing the increased inhibition of losing competitors. Thicker lines indicate stronger excitatory or inhibitory signals. See text for further explanation.

Action selection in the BG can improve through learning. The BG are strongly implicated in instrumental conditioning and in various forms of sequential learning. An important group of neurons in the midbrain BG areas appear to fire in conjunction with rewarding events, or prior to anticipated rewarding events. Shultz, Dayan, and Montague (1997) have proposed that the afferents from these structures to the striatum may provide a training signal similar to the *temporal difference error* used in artificial reinforcement learning methods (see e.g. Barto, 1995). Houk, Adams and Barto (1995) have further speculated that something akin to an *actor-critic* learning system (Barto, 1995) may be operating in the BG. Although these comparisons have yet to be fully substantiated, there is good evidence that these circuits could support adaptive tuning of selection mechanisms.

Evolutionary considerations. The fundamental architecture of the basal ganglia is archaic in evolutionary terms. We have already pointed out that the main basal ganglia nuclei have homologues in all the vertebrate groups, it further appears that the neurotransmitter organization (cell type) and connectivity patterns of the basal ganglia are largely conserved in at least the jawed vertebrates (Medina & Reiner, 1995), and that the striatum occupies a roughly similar proportion of forebrain volume in all vertebrate classes (Hodos, 1982). This evidence suggests that the BG may be part of the basic vertebrate brain plan. The limited neurobehavioral data available for reptiles (Tarr, 1982) and amphibians (Ewert, 1995) adds further support to the notion of a core BG function that is similar in all vertebrates.

An important characteristic of the non-mammalian BG is that the primary output pathways are directed to structures in the midbrain (Medina & Reiner, 1995). This would suggest that the original role of the BG was to arbitrate between the different demands of brainstem sensorimotor systems. With the evolution of corticospinal motor pathways in the mammalian brain, selection circuits in the BG may have been recruited to serve a similar role with respect to the newer cortical motor systems.

The Granularity of Selection in the Basal Ganglia. An important issue concerns the granularity of the action selection subserved by the BG. For instance, some researchers have proposed a role for the BG in a more fine-grained sequencing of movement than selecting between competing behavioral acts (see Mink, 1996, for review). The suggestion that the BG is involved in movement sequencing can, however, be reconciled with the view of the BG as an action selection device on the grounds that both tasks can be regarded as similar problems on somewhat different time-scales. We propose a general hypothesis which appears to be consistent with the anatomy of the BG, namely, that similar switching circuitry is employed in different regions of the BG to resolve selection problems at different levels of functional integration. Specifically, the BG is widely described as having three main territories termed the *limbic, associative*, and *motor* domains (Gerfen & Wilson, 1996; Joel & Weiner, 1994). These areas have similar intrinsic circuitry but different extrinsic connectivity. There is evidence that the limbic striatum (also termed the nucleus accumbens) receives connections from many areas of the brain that are involved in generating motivated behavior (see Groenewegen et al., 1996). It seems possible, therefore, that competitions at higher levels of functional integration are resolved in the limbic domain of the BG. At the other end of the scale, the motor domain appears to be specialized to resolve conflicts for specific muscle groups, suggesting a role in selection at a comparatively low levels of functional integration—choosing between incompatible motor patterns. The role of the associative domain is less clearly characterized by existing research, however, a possibility arising from the current theory is that it plays a role in selection at intermediate levels, for instance, in choosing between co-ordinated movement sequences. Joel and Weiner (1994) provide a useful review of some of the neural circuitry that might mediate interactions between the different striatal domains. This evidence suggests that such interactions are likely to be indirect (i.e. involve structures outside the BG), however, they will clearly be important in co-ordinating appropriate patterns of selection at multiple spatial and temporal scales.

The Brain Without the Basal Ganglia. Further insight into the role of the BG within the overall control architecture of the vertebrate brain can be gained by considering circumstances in which BG circuitry is either disconnected or inoperative. For instance, a number of studies have looked at the effect of blocking the outputs from the basal ganglia to the rat superior colliculus (see Dean et al., 1989). From our perspective this can viewed as jamming the switching circuit into the "on" position (no BG inhibition of collicular output systems). In these circumstances the animal has no difficulty in co-ordinating the head and mouth movements required to locate and bite an object moving in its whisker field, however, the animal also appears unable to resist orienting to any tactile stimulus placed in the whisker field and fails to habituate to such a stimulus (Dean et al., 1989; Redgrave, personal observation). This inability to suppress an activated, but non-salient, sensorimotor system indicates a dissociation between the circuitry which controls switching (in the BG), and that which controls planning and execution of the motor act (in the colliculus and brainstem). Other research has looked at the effect of jamming the BG in the "off" position. For instance, Teitelbaum and co-workers have performed extensive observational studies using animals with lesions of the lateral hypothalmus (reviewed in Teitelbaum et al, 1983; Teitelbaum, Pellis & Pellis, 1990). Such lesions destroy the dopaminergic inputs to the striatum, rendering the disinhibitory BG pathway inoperative, and resulting in continuous inhibition of all motor systems. Experimental animals show no spontaneous, environmentally-oriented movement for a day or so after the lesion operation and must be tube-fed to stay alive. With time, however, the ability to perform co-ordinated movements recovers to the extent that the animal will walk around, explore, and eat any palatable food it encounters. Behavior-generating mechanisms in the brain thus appear to be able to adapt so as to compensate for malfunction in the BG switching circuitry. What fails to recover, however, is the natural ease with which intact animals organize their actions into coherent motivated sequences. For example, recovered animals, whose behavior appears relatively normal, will often walk into a corner and stand for a long time making stereotyped stepping and head-scanning movements. Teitelbaum et. al. term this situation a "behavioral trap" and describe the animal's behavior as "disintegrated" and "disconnected from its usual adaptive outcome" (Teitelbaum et al., 1990 p. 101). Such observations may strike a chord with researchers in adaptive behavior who have witnessed a behavior-based robot, trapped in the corner of a room, repetitiously alternating between left and right avoidance reactions. This suggests that the design of future robot architectures might benefit from understanding how similar problems are resolved in the vertebrate BG.

CONCLUSIONS

This article has reviewed the architecture and evolution of the vertebrate nervous system and drawn comparisons with the subsumption architecture (SA) proposed by Brooks' as a design methodology for building mobile robots. We have argued that an important dimension in the organization of the vertebrate brain is a division into multiple, vertical layers, and we have suggested, using the example of the rat defense system, that the subsumption architecture provides a useful metaphor for understanding many of the interactions between these layers.

Vertebrate behavior is, however, determined by multiple inter-twined behavior systems that can be in conflict on several levels. We have argued that some of these conflicts cannot be easily resolved through fixed, subsumption-like links or through emergent switching processes. In particular, we have suggested that conflicts that arise between functional units that are widely distributed in the brain may be resolved by centrally-located, specialized selection mechanisms in the basal ganglia. This suggestion could provide a way to reconcile the Jacksonian view of the brain as a layered control system with the idea of a centrencephalic dimension to brain organization.

The notion of incremental evolution employed in the subsumption architecture was considered to capture some important characteristics of the evolution of the vertebrate nervous system; for example, that neural centers tend to retain much of their function as newer circuits evolve alongside, and that evolution often prefers to conserve and suppress redundant neural circuits rather than excise them. We also noted, however, the remarkable conservation of the basic vertebrate brain plan, which has survived adaptation and divergence to life in a wide range of adaptive zones. If the success of the vertebrate sub-phylum is in part due to this generic brain architecture, it seems likely that many of its characteristics could be usefully copied in the control systems of autonomous robots.

ACKNOWLEDGMENTS

This research was partially supported by Wellcome trust grant 038011/Z/93 to Peter Redgrave. We are grateful to Paul Dean, Max Westby, and John Mayhew, for discussion of the subject matter of this article.

REFERENCES

Albus, J. S. (1991). Outline for a theory of intelligence. *IEEE Transactions On Systems Man and Cybernetics*, **21**, 473-509.

Altman, J. S. & Kien, J. (1989). New models of motor control. *Neural Computation,* **1**, 173-183.

Arbib, M. A. (1989). *The metaphorical brain 2: Neural networks and beyond.* New York: Wiley.

Arbib, M. A. & Liaw, J. S. (1995). Sensorimotor transformations in the worlds of frogs and robots. *Artificial Intelligence,* **72**, 53-79.

Baerends, G. P. (1970). A model of the functional organization of incubation behaviour. *Behaviour Supplement,* **17**, 263-312.

Baerends, G. P. (1976). The functional organization of behaviour. *Animal Behaviour,* **24**, 726-738.

Bakshi, V. P. & Kelley, A. E. (1991). Dopaminergic regulation of feeding-behavior .1. Differential effects of haloperidol microinfusion into 3 striatal subregions. *Psychobiology,* **19**, 223-232.

Ballard, D. H., Hayhoe, M. M. & Whitehead, S. D. (1992). Hand-eye coordination during sequential tasks. *Philosophical Transactions of the Royal Society of London. B,* **337**, 331-339.

Barto, A. G. (1995). Adaptive critics and the basal ganglia. In J. C. Houk, J. L. Davis & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 215–232). Cambridge, MA: MIT Press.

Belekhova, M. G. (1990). Limbic system and the problem of evolution in the telencephalon of vertebrates. *Journal of Evolutionary Biochemistry and Physiology,* **26**, 411-422.

Belekhova, M. G. & Veselkin, N. P. (1985). Telencephalization and transfer of function in the central nervous system of vertebrates in light of current data. *Journal of Evolutionary Biochemistry and Physiology,* **21**, 357-365.

Berntson, G. G., Boysen, S. T. & Cacioppo, J. T. (1993). Neurobehavioral organization and the cardinal principle of evaluative bivalence. *Annals of the New York Academy of Sciences,* **702**, 75-102.

Bizzi, E., Giszter, S. F., Loeb, E., Mussaivaldi, F. A. & Saltiel, P. (1995). Modular organization of motor behavior in the frog's spinal-cord. *Trends In Neurosciences,* **18**, 442-446.

Blanchard, R. J. & Blanchard, D. C. (1987). An ethoexperimental approach to the study of fear. *The Psychological Record,* **37**, 305-316.

Blumberg, B. (1994). Action selection in Hamsterdam: lessons from ethology. In D. Cliff, P. Husbands, J.-A. Meyer & S. W. Wilson (Eds.), *From animals to animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior* (pp. 22-29). Cambridge, MA: MIT Press.

Braitenberg, V. (1986). *Vehicles: experiments in synthetic psychology.* Cambridge, MA: MIT Press.

Brandao, M. L., Melo, L. L. & Cardoso, S. H. (1993). Mechanisms of defense in the inferior colliculus. *Behavioural Brain Research,* **58**, 49-55.

Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal on Robotics and Automation,* **2**, 14-23.

Brooks, R. A. (1989). A robot that walks: emergent behaviour from a carefully evolved network. *Neural Computation,* **1**, 253-262.

Brooks, R. A. (1990). Elephants don't play chess. In P. Maes (Ed.), *Designing autonomous agents: Theory and practice from biology to engineering and back* (pp. 3–15). Cambridge, MA: MIT Press.

Brooks, R. A. (1991a). New approaches to robotics. *Science,* **253**, 1227-1232.

Brooks, R. A. (1991b). Intelligence without representation. *Artificial Intelligence,* **47**, 139-159.

Brooks, R. A. (1994). Coherent behaviour from many adaptive processes. In D. Cliff, P. Husbands, J.-A. Meyer & S. W. Wilson (Eds.), *From animals to animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior* (pp. 22–29). Cambridge, MA: MIT Press.

Brooks, R. A. (1995). Intelligence without reason. In L. Steels & R. Brooks (Eds.), *The artificial life route to artificial intelligence* (pp. 25–81). Hillsdale, New Jersey: Lawrence Erlbaum.

Bruce, L. L. & Neary, T. J. (1995). The limbic system of tetrapods—a comparative analysis of cortical and amygdalar populations. *Brain, Behavior, and Evolution,* **46**, 224-234.

Butler, A. B. & Hodos, W. (1996). *Comparative vertebrate neuroanatomy.* New York: Wiley-Liss.

Casseday, J. H. & Covey, E. (1996). A neuroethological theory of the operation of the inferior colliculus. *Brain, Behavior, and Evolution,* **47**, 311-336.

Chapman, D. (1991). *Vision, instruction, and action.* Cambridge, MA: MIT Press.

Chevalier, G. & Deniau, J. M. (1990). Disinhibition as a basic process in the expression of striatal functions. *Trends in Neurosciences,* **13**, 277-280.

Cho, J. W. & West, M. O. (1997). Distributions of single neurons related to body parts in the lateral striatum of the rat. *Brain Research,* **756**, 241-246.

Cliff, D. T. (1991). Computational neuroethology: a provisional manifesto. In J.-A. Meyer & S. Wilson (Eds.), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior* (pp. 29–39). Cambridge, MA: MIT Press.

Cohen, A. H. (1988). Evolution of the vertebrate central pattern generator for locomotion. In A. H. Cohen, S. Rossignol & S. Grillner (Eds.), *Neural control of rhythmic movements in vertebrates* (pp. 129–166). New York: Wiley.

Cohen, A. H. (1992). The role of heterarchical control in the evolution of central pattern generators. *Brain, Behavior, and Evolution,* **40**, 112-124.

Colgan, P. (1989). *Animal motivation.* London: Chapman & Hall.

Connell, J. H. (1990). *Minimalist mobile robots.* Boston: Academic Press.

Davey, G. (1989). *Ecological learning theory.* London: Routledge.

Davis, M. (1984). The mammalian startle response. In R. C. Eaton (Ed.), *Neural mechanisms of startle behaviour* (pp. 287–351). New York: Plenum Publishing Co.

Davis, M. (1992). The role of the amygdala in fear and anxiety. *Annual Review of Neuroscience,* **15**, 353-75.

Deacon, T. W. (1990). Rethinking mammalian brain evolution. *American Zoologist,* **30**, 629-705.

Dean, P., Redgrave, P. & Westby, G. W. M. (1989). Event or emergency—2 response systems in the mammalian superior colliculus. *Trends In Neurosciences,* **12**, 137-147.

Depaulis, A., Keay, K. A. & Bandler, R. (1992). Longitudinal neuronal organization of defensive reactions in the midbrain periaqueductal gray region of the rat. *Experimental Brain Research,* **90**, 307-318.

Edwards, D. H. (1991). Mutual inhibition among neural command systems as a possible mechanism for behavioral choice in crayfish. *Journal of Neuroscience,* **11**, 1210-1223.

Ewert, J-P. (1995). Command neurons and command systems. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 215–220). Cambridge, MA: MIT Press.

Fanselow, M. S. & Sigmundi, R. A. (1987). Functional behaviourism and aversively motivated behavior: a role for endogenous opioids in the defense system of the rat. *Psychological Record,* **37**, 317-334.

Fetcho, J. R. (1992). The spinal motor system in early vertebrates and some of its evolutionary changes. *Brain, Behaviour, and Evolution,* **40**, 82-97.

Finlay, B. L., Wikler, K. C. & Sengelaub, D. R. (1987). Regressive events in brain-development and scenarios for vertebrate brain evolution. *Brain, Behavior, and Evolution,* **30**, 102-117.

Franceschini, N. (1992). From insect vision to robot vision. *Philosophical Transactions of the Royal Society of London. B,* **337**, 283-294.

Fritzsch, B. (1995). Evolution of the ancestral vertebrate brain. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 373–377). Cambridge, MA: MIT Press.

Gallistel, C. R. (1980). *The organization of action: a new synthesis.* Hillsdale, NJ: Lawrence Erlbaum.

Gans, C. (1989). Stages in the origin of vertebrates—analysis by means of scenarios. *Biological Reviews of the Cambridge Philosophical Society,* **64**, 221-268.

Gans, C. & Northcutt, R. G. (1983). Neural crest and the origin of vertebrates—a new head. *Science,* **220**, 268-274.

Gelissen, M. & Cools, A. (1988). Effect of intracaudate haloperidol and apomorphine on switching motor patterns upon current behavior of cats. *Behavioural Brain Research,* **29**, 17-26.

Gerfen, C. R. & Wilson, C. J. (1996). The basal ganglia. In L. W. Swanson, A. Bjorklund and T. Hokfelt (Eds.), *Handbook of chemical neuroanatomy, Vol 12: Integrated systems of the CNS, Part III* (pp. 371–468). Amsterdam: Elsevier.

Getting, P. A. (1989). Emerging principles governing the operation of neural networks. *Annual Review of Neuroscience,* **12**, 185-204.

Gurney, K., Prescott, T. J., & Redgrave, P. (1998). The basal ganglia viewed as an action selection device. In Niklasson, L., Bodén, M. & Ziemke, T. (Eds.), *Proceedings of the Eighth International Conference on Artificial Neural Networks* (pp. 1033-1038). Berlin: Springer.

Groenewegen, H. J., Wright, C. I. & Beijer, A. V. J. (1996). The nucleus accumbens: Gateway for limbic structures to reach the motor system? *Progress in Brain Research,* **107**, 485-511.

Harris-Warrick, R. M. (1991). Modulation of neural networks for behaviour. *Annual Review of Neuroscience,* **14**, 39-57.

Harvey, P. H. & Krebs, J. R. (1990). Comparing brains. *Science,* **249**, 140-146.

Hodos, W. (1982). Some perspectives on the evolution of intelligence and the brain. In D. R. Griffin (Ed.), *Animal mind—human mind* (pp. 33–56). Berlin: Springer-Verlag.

Holland, P. W. H. & Graham, A. (1995). Evolution of regional identity in the vertebrate nervous-system. *Perspectives On Developmental Neurobiology,* **3**, 17-27.

Houk, J. C., Adams, J. L. & Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. C. Houk, J. L. Davis & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 249–270). Cambridge, MA: MIT Press.

Houston, A. & Sumida, B. (1985). A positive feedback model for switching between 2 activities. *Animal Behaviour,* **33**, 315-325.

Humphrey, D. R. (1986). Representation of movements and muscles within the primate precentral motor cortex: historical and current perspectives. *Federation Proceedings,* **45**, 2687-2699.

Jackson, J. H. (1958). Evolution and dissolution of the nervous system. In J. Taylor (Ed.), *Selected writings of John Hughlings Jackson* (Vol. 2, pp. 45–75). London: Staples Press. (Original work published 1884).

Jerison, H. (1973). *Evolution of the brain and intelligence.* New York: Academic Press.

Joel, D. & Weiner, I. (1994). The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. *Neuroscience,* **63**, 363-379.

Kavanau, J. L. (1990). Conservative behavioral evolution, the neural substrate. *Animal Behaviour,* **39**, 758-767.

Kelso, J. A. S. (1995). *Dynamic Patterns: The Self-organization of Brain and Behaviour.* Cambridge, MA: MIT Press.

Kien, J. (1986). From command fibers to command systems to consensus—are these labels really useful anymore. *Behavioral and Brain Sciences,* **9**, 732-733.

King, J. R. & Comer, C. M. (1996). Visually elicited turning behavior in rana-pipiens—comparative organization and neural control of escape and prey capture. *Journal Of Comparative Physiology A,* **178**, 293-305.

King, S. M., Dykeman, C., Redgrave, P. & Dean, P. (1992). Use of a distracting task to obtain defensive head movements to looming visual stimuli by human adults in a laboratory setting. *Perception,* **21**, 245-259.

Lacalli, T. C. (1996). Frontal eye circuitry, rostral sensory pathways and brain organization in Amphioxus larvae—evidence from 3d reconstructions. *Philosophical Transactions Of the Royal Society Of London Series B,* **351**, 243-263.

Ledoux, J. E. (1992). Emotion and the amygdala. In J. P. Aggleton (Ed.), *The amygdala: Neurological aspects of emotion, memory, and mental dysfunction* (pp. 339–351). New York: Wiley-Liss.

LeDoux, J. E. (1995). Emotion: clues from the brain. *Annual Review of Psychology,* **46**, 206-235.

Leise, E. M. (1990). Modular construction of nervous systems: a basic principle of design for invertebrates and vertebrates. *Brain Research Reviews,* **15**, 1-23.

Lingenhohl, K. & Friauf, E. (1994). Giant-neurons in the rat reticular-formation—a sensorimotor interface in the elementary acoustic startle circuit. *Journal of Neuroscience,* **14**, 1176-1194.

Ludlow, A. R. (1976). The behaviour of a model animal. *Behaviour,* **58**, 131-172.

Maes, P. (1990). Situated agents can have goals. *Robotics and Autonomous Systems,* **6**, 49-70.

Maes, P. (1992). Behaviour-based artificial intelligence. In J.-A. Meyer, H. L. Roitblat & S. W. Wilson (Eds.), *From animals to animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior* (pp. 238–246). Cambridge, MA: MIT Press.

Maes, P. (1995). Modeling adaptive autonomous agents. In C. G. Langton (Ed.), *Artificial life: an overview* (pp. 135–162). Cambridge, MA: MIT Press.

Magoun, H. W. (1958). *The waking brain.* Springfield, Illinois: Thomas.

McFarland, D. (1965). Flow graph representation of motivational systems. *British Journal of Mathematical and Statistical Psychology,* **18**, 25-43.

McFarland, D. (1971). *Feedback mechanisms in animal behaviour.* London: Academic Press.

McFarland, D. (1989). *Problems of animal behaviour.* Harlow, UK: Longman.

McFarland, D. & Spier, E. (1997). Basic cycles, utility and opportunism in self-sufficient robots. *Robotics and Autonomous Systems,* **20**, 179-90.

McGonigle, B. (1990). Incrementing intelligent systems by design. In J.-A. Meyer & S. Wilson (Eds.), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior* (pp. 525–531). Cambridge, MA: MIT Press.

Medina, L. & Reiner, A. (1995). Neurotransmitter organization and connectivity of the basal ganglia in vertebrates: implications for the evolution of the basal ganglia. *Brain, Behavior, and Evolution,* **46**, 235-258.

Meredith, M. A. & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science,* **221**, 220-222.

Metcalfe, W. K., Mendelson, B. & Kimmel, C. B. (1986). Segmental homologies among reticulospinal neurons in the hindbrain of the zebrafish larva. *Journal of Comparative Neurology,* **251**, 147-159.

Meyer, J. & Guillot, A. (1990). Simulation of adaptive behaviour in animats: review and prospect. In J.-A. Meyer & S. Wilson (Eds.), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior* (pp. 2–11). Cambridge, MA: MIT Press.

Miklos, G. L. G. (1993). Molecules and cognition: the latterday lessons of levels, language, and iac. *Journal of Neurobiology,* **24**, 842-890.

Mink, J. W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. *Progress In Neurobiology,* **50**, 381-425.

Minsky, M. L. (1986). *The Society of Mind.* New York: Simon & Schuster.

Nauta, W. J. H. & Karten, H. J. (1970). A general profile of the vertebrate brain, with sidelights on the ancestry of the cerebral cortex. In F. O. Schmitt (Ed.), *The neurosciences: Second study program* (pp. 7–26). New York: Rockefeller University Press.

Northcutt, R. G. (1981). Evolution of the telencephalon in nonmammals. *Annual Review of Neuroscience,* **4**, 301-350.

Northcutt, R. G. (1984). Evolution of the vertebrate central nervous-system—patterns and processes. *American Zoologist,* **24**, 701-716.

Northcutt, R. G. (1994). An immunohistochemical study of the telencephalon and the diencephalon in a myxinoid jawless fish, the pacific hagfish, Eptatretus stouti. *Brain, Behavior, and Evolution,* **43**, 140-161.

Northcutt, R. G. (1996). The agnathan ark—the origin of craniate brains. *Brain, Behavior, and Evolution,* **48**, 237-247.

Northcutt, R. G. and Kaas, J. H. (1995). The emergence and evolution of mammalian neocortex. *Trends In Neurosciences,* **18**, 373-379.

Overmier, J. B. & Hollis, K. L. (1990). Fish in the think tank: Learning, memory, and integrated behaviour. In R. P. Kesner & D. S. Olton (Eds.), *Neurobiology of comparative cognition* (pp. 205–236). Hillsdale, NJ: Lawrence Erlbaum.

Penfield, W. (1975). *The mystery of the mind.* Princeton, NJ: Princeton University Press.

Prescott, T. J. (1995). Spatial representation for navigation in animats. *Adaptive Behavior,* **4**, 85-123.

Redgrave, P. & Dean, P. (1991). Does the PAG learn about emergencies from the superior colliculus ? In A. Depaulis & R. Bandler (Eds.), *The midbrain periaqueductal gray matter* (pp. 199–209). New York: Plenum Press.

Redgrave, P., Prescott, T. J., & Gurney, K. (In press). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience*.

Reiner, A. (1994). Catecholaminergic innervation of the basal ganglia in mammals: anatomy and function. In W. J. A. J. Smeets & A. Reiner (Eds.), *Phylogeny and Development of Catecholamine Systems in the CNS of Vertebrates* (pp. 247–272). Cambridge, UK: Cambridge University Press.

Ringo, J. L. (1991). Neuronal interconnection as a function of brain size. *Brain, Behaviour, and Evolution,* **38**, 1-6.

Roitblat, H. L. (1991). Cognitive action theory as a control architecture. In J.-A. Meyer & S. Wilson (Eds.), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior* (pp. 444–450). Cambridge, MA: MIT Press.

Rosenblatt, J. K. & Payton, D. W. (1989). A fine-grained alternative to the subsumption architecture for mobile robot control. In *Proceedings of the IEEE/INNS International Joint Conference on Neural Networks* (Vol. 2, pp. 317–324).

Rosenblatt, J. K. (1997). The distributed architecture for mobile navigation. *Journal of Experimental and Theoretical Artificial Intelligence*, **9**, 339-360.

Rudy, J. W., Stadler-Morris, S. & Albert, P. (1987). Ontogeny of spatial navigation behaviours in the rat: dissociation of "proximal"- and "distal"- cue-based behaviours. *Behavioural Neuroscience,* **101**, 62-73.

Rutkowska, J. C. (1994). Scaling up sensorimotor systems: constraints from human infancy. *Adaptive Behavior,* **2**, 349-373.

Salamone, J. D. (1991). Behavioral pharmacology of dopamine systems: a new synthesis. In P. Willner & J. Scheel-Kruger (Eds.), *The mesolimbic dopamine system: From motivation to action* (pp. 599–613). New York: Wiley.

Salamone, J. D., Cousins, M. S. & Snyder, B. J. (1997). Behavioral functions of nucleus accumbens dopamine: Empirical and conceptual problems with the anhedonia hypothesis. *Neuroscience and Biobehavioral Reviews,* **21**, 341-359.

Sarnat, H. B. & Netsky, M. G. (1981). *Evolution of the nervous system* (2nd ed.). Oxford: Oxford University Press.

Schoner, G., Dose, M. & Engels, C. (1995). Dynamics of behavior: theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems,* **16**, 213-245.

Schultz, W., Dayan, P. & Montague, P. R. (1997). A neural substrate for prediction and reward. *Science,* **275**, 1593-1599.

Sherrington, C. S. (1947). *The integrative action of the nervous system* (Rev. Ed.). New Haven: Yale University Press. (1st edition 1906).

Snaith, S. & Holland, O. (1990). An investigation of two mediation strategies suitable for behavioural control in animals and animats. In J.-A. Meyer & S. Wilson (Eds.), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior* (pp. 255–262). Cambridge, MA: MIT Press.

Steels, L. (1995). The artificial life roots of artificial intelligence. In C. G. Langton (Ed.), *Artificial life: an overview* (pp. 75–110). Cambridge, MA: MIT Press.

Stein, B. E. & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.

Szentagothai, J. & Arbib, M. A. (1975). *Conceptual models of neural organization*. Cambridge, MA: MIT Press.

Tarr, R. S. (1982). Species typical display behavior following stimulation of the reptilian striatum. *Physiology and Behavior,* **29**, 615-620.

Teitelbaum, P., Pellis, V. C. & Pellis, S. M. (1990). Can allied reflexes promote the integration of a robot's behavior? In J.-A. Meyer & S. Wilson (Eds.), *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior* (pp. 97–104). Cambridge, MA: MIT Press.

Teitelbaum, P., Schallert, T. & Whishaw, I. Q. (1983). Sources of spontaneity in motivated behaviour. In P. Teitelbaum & E. Satinoff (Eds.), *Handbook of behavioural neurobiology.* (Vol. 6, pp. 23–65). New York: Plenum Press.

Thompson, R. (1993). Centrencephalic theory, the general learning system, and subcortical dementia. *Annals of the New York Academy of Sciences,* **702**, 197-223.

Timberlake, W. (1993). Animal behavior: a continuing synthesis. *Annual Review of Psychology,* **44**, 675-708.

Tyrrell, T. (1992). Defining the action selection problem. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 1152–1157). Hillsdale, NJ: Lawrence Erlbaum.

Tyrrell, T. (1993). The use of hierarchies for action selection. *Adaptive Behavior,* **1**, 387-420.

Van Sommers, P. (1972). *The biology of behaviour.* New York: Wiley.

Wagner, G. P. & Altenberg, L. (1996). Perspective—complex adaptations and the evolution of evolvability. *Evolution,* **50**, 967-976.

Walker, S. (1983). *Animal thought.* London: Routledge & Kegan Paul.

Wallace, M. T. & Stein, B. E. (1997). Development of multisensory neurons and multisensory integration in cat superior colliculus. *Journal of Neuroscience,* **17**, 2429-2444.

Walter, W. G. (1953). *The living brain.* London: Duckworth.

Wickens, J. (1997). Basal ganglia: structure and computations. *Network—Computation in Neural Systems,* **8**(4), R77-R109.

Wilczynski, W. (1984). Central neural systems subserving a homoplasous periphery. *American Zoologist,* **24**, 755-763.

Windhorst, U. (1996). On the role of recurrent inhibitory feedback in motor control. *Progress In Neurobiology,* **49**, 517-587.

Wise, S. P. (1993). Monkey motor cortex: movements, muscles, motoneurons and metrics. *Trends in Neuroscience,* **16**, 46-49.

Whishaw, I. Q. (1990). The decorticate rat. In B. Kolb & R. C. Tees (Eds.), *The cerebral cortex of the rat.* Cambridge, MA: MIT Press.

Wurtz, R. H. & Hikosaka, O. (1986). Role of the basal ganglia in the initiation of saccadic eye movements. *Progress in Brain Research,* **64**, 175-189.

Yardley, C. P. & Hilton, S. M. (1986). The hypothalamic and brainstem areas from which the cardiovascular and behavioural components of the defence reaction are elicited in the rat. *Journal of the Autonomic Nervous System,* **15**, 227-244.