Cortical mechanisms of action selection: the affordance competition hypothesis

Paul Cisek*

Department of physiology, University of Montréal, C.P. 6128 Succursale Centre-ville, Montréal, Quebec, H3C 3J7 Canada

At every moment, the natural world presents animals with two fundamental pragmatic problems: selection between actions that are currently possible and specification of the parameters or metrics of those actions. It is commonly suggested that the brain addresses these by first constructing representations of the world on which to build knowledge and make a decision, and then by computing and executing an action plan. However, neurophysiological data argue against this serial viewpoint. In contrast, it is proposed here that the brain processes sensory information to specify, in parallel, several potential actions that are currently available. These potential actions compete against each other for further processing, while information is collected to bias this competition until a single response is selected. The hypothesis suggests that the dorsal visual system specifies actions which compete against each other within the fronto-parietal cortex, while a variety of biasing influences are provided by prefrontal regions and the basal ganglia. A computational model is described, which illustrates how this competition may take place in the cerebral cortex. Simulations of the model capture qualitative features of neurophysiological data and reproduce various behavioural phenomena.

Keywords: decision making; action selection; cerebral cortex; computational modelling

1. INTRODUCTION

At every moment, the natural environment presents animals with many opportunities and demands for action. The presence of food offers an opportunity to satiate hunger, while the appearance of a predator demands caution or evasion. An animal cannot perform all these behaviours at the same time because they often share the same effectors (you only have two hands; you can only transport yourself in one direction at a time, etc.). Thus, one fundamental issue faced by every behaving creature is the question of action selection. This question must be resolved, in part, by using external sensory information about objects in the world and, in part, by using internal information about the current behavioural needs.

Furthermore, the animal must tailor the actions it performs to the environment in which it is situated. Grasping a fruit requires accurate guidance of the hand to the location of the fruit, while evading a predator requires one to move in an unobstructed direction that leads away from the threat. The specification of the parameters of actions is a second fundamental issue faced by behaving creatures. Specification of actions also must use sensory information from the environment. In particular, it requires information about the spatial relationships among objects and surfaces in the world, represented in a coordinate frame relative to the orientation and configuration of the animal’s body.

Traditional cognitive theories propose that these two questions are resolved in a serial manner, so that we select ‘what to do’ before specifying ‘how to do it’. According to this view, the perceptual system first collects sensory information to build an internal descriptive representation of objects in the external world (Marr 1982). Next, this information is used along with representations of current needs and memories of past experience to make judgments and decide upon a course of action (Newell & Simon 1972; Johnson-Laird 1988; Shafir & Tversky 1995). The resulting plan is then used to generate a desired trajectory for movement, which is finally realized through muscular contraction (Miller et al. 1960; Keele 1968). In other words, the brain first builds knowledge about the world using representations which are independent of actions, and this knowledge is later used to make decisions, compute an action plan and finally execute a movement.

However, studies on the cerebral cortex have encountered difficulties in interpreting neural activity in terms of distinct perceptual, cognitive or motor systems. For example, visual processing diverges in the cortex into separate systems sensitive to object identity and spatial location (Ungerleider & Mishkin 1982), with no single representation of the world (Stein 1992), leading to the question of how these disparate systems are bound together to form a unified percept (von der Malsburg 1996; Cisek & Turgeon 1999). Cells in the posterior parietal cortex appear to reflect a mixture of sensory (Andersen 1995; Colby & Goldberg 1999), motor (Snyder et al. 1997) and cognitive information (Platt & Glimcher 1999), leading to persistent debates on their functional role. A recent review of data on the parietal cortex has suggested that ‘current hypotheses concerning...’
parietal function may not be the actual dimensions along which the parietal lobes are functionally organized; on this view, what we are lacking is a conceptual advance that leads us to test better hypotheses” (Culham & Kanwisher 2001, pp. 159–160). In other words, perhaps the concepts of separate perceptual, cognitive and motor systems, which theoretical neuroscience inherits from cognitive psychology, are not appropriate for bridging neural data with behaviour.

Even stronger concerns regarding cognitive psychology’s suitability as a bridging framework are raised by considerations of evolutionary history (Sterelny 1989; Hendriks-Jansen 1996). Brain evolution is strikingly conservative and major features of modern neural organization can be seen in the humble Haikouichthys, a primitive jawless fish that lived during the Early Cambrian epoch over 520 Myr ago (Shu et al. 2003). Since the development of the telencephalon, the basic outline of the vertebrate nervous system has been strongly conserved throughout its phylogenetic history (Butler & Hodos 1996; Holland & Holland 1999; Katz & Harris-Warrick 1999) and even recently, elaborated structures such as the mammalian neocortex have homologues among non-mammalian species (Medina & Reiner 2000). Although the idea that brain evolution consists of new structures being added on top of old structures (e.g. the ‘Triune Brain’; MacLean 1973) is still popular among non-specialists, it has been rejected in recent decades of comparative neuroanatomical work (Deacon 1990; Butler & Hodos 1996). Brain evolution consists of the differentiation and specialization of existing structures through shifts in existing axonal projection patterns (Deacon 1990; Krubitzer & Kaas 2005), not through the addition of new structures. Thus, the basic anatomical and functional organization of the primate brain reflects an ancient architecture which was well established by the time of the earliest terrestrial tetrapods. This architecture could not have been designed to serve the needs of higher cognitive abilities, which did not exist, but must have been laid down so as to best address the needs of simple interactive behaviour.

An emphasis on the brain’s role in interactive behaviour is by no means novel. Similar ideas have for a long time been central to theories in ethology (Hinde 1966; Ewert et al. 2001) and have recently led to several viewpoints on cognition (Adams & Mele 1989; Clark 1997; Beer 2000; Núñez & Freeman 2000; Thelen et al. 2001) and interactive behaviour (Brooks 1991; Hendriks-Jansen 1996; Prescott et al. 1999; Seth 2007). All these are similar to several lines of thought that are much older (Mead 1938; Merleau-Ponty 1945; Ashby 1965; Powers 1973; Gibson 1979; Maturana & Varela 1980) in some cases by over a hundred years (Jackson 1884; Bergson 1896; Dewey 1896). Most of these viewpoints emphasize the pragmatic aspects of behaviour (Piaget 1967; Gibson 1979; Millikan 1989), a theme that underlies several proposals regarding representation (Dretske 1981; Gallese 2000; Hommel et al. 2001), memory (Ballard et al. 1995; Glennberg 1997) and visual consciousness (O’Regan & Noé 2001). Here, it is proposed that these views, which emphasize the brain’s role in controlling behaviour in real time (Cisek 1999), provide a better basis for interpreting neurophysiological data than the traditional framework of cognitive psychology (Cisek 2001).

Continuous interaction with the world often does not allow one to stop to think or to collect information and build a complete knowledge of one’s surroundings. To survive in a hostile environment, one must be ready to act at short notice, releasing into execution actions that are at least partially prepared. These are the fundamental demands which shaped brain evolution. They motivate animals to process sensory information in an action-dependent manner to build representations of the potential actions which the environment currently affords. In other words, the perception of a given natural setting may involve not only representations which capture information about the identity of objects in the setting, but also representations which specify the parameters of possible actions that can be taken (Gibson 1979; Fadiga et al. 2000; Cisek 2001). With a set of such potential actions partially specified, the animal is ready to quickly perform actions if circumstances demand. In essence, it is possible that the nervous system addresses the questions of specification (how to do it) before performing selection (what to do). Indeed, for continuous interactive behaviour, it may be best to perform both specification and selection processes at all times to enable continuous adjustment to the changing world.

The proposal made here is that the processes of action selection and specification occur simultaneously and continue even during overt performance of movements. That is, sensory information arriving from the world is continuously used to specify several currently available potential actions, while other kinds of information are collected to select from among these the one that will be released into overt execution at a given moment (Kalaska et al. 1998; Kim & Shadlen 1999; Cisek 2001; Glimcher 2001; Gold & Shadlen 2001; Platt 2002; Cisek & Kalaska 2005). From this perspective, behaviour is viewed as a constant competition between internal representations of the potential actions which Gibson (1979) termed ‘affordances’. Hence, the framework presented here is called the ‘affordance competition hypothesis’.

It is not proposed that complete action plans are prepared for all of the possible actions that one might take at a given moment. First, only actions which are currently available are specified in this manner. Second, many possible actions are eliminated from processing by selective attention mechanisms which limit the sensory information that is transformed into representations of action. Finally, complete action planning is not proposed even for the final selected action. Even in cases of highly practised behaviours, no complete pre-planned motor programme or entire desired trajectory appears to be prepared (Kalaska et al. 1998; Cisek 2005).

2. THE AFFORDANCE COMPETITION HYPOTHESIS

The view of behaviour as a competition between actions has been common in studies on animal behaviour and the interpretation of subcortical circuits (Ewert 1997; Prescott et al. 1999; Ewert et al. 2001). However, it is more rarely used to explain the activity of...
cerebral cortical regions, perhaps, owing to an assumption that the cortex is a new structure concerned with new cognitive functions. However, as discussed above, this assumption is not justified. The organization of the cerebral cortex has been conserved for a long time, motivating one to interpret it, like subcortical circuits, in terms of interactive behaviour. Figure 1 outlines a proposal on how the affordance competition hypothesis may be used to interpret neural data from the primate cerebral cortex during visually guided behaviour.

The visual system is organized into two parallel processing pathways: an occipito-temporal ‘ventral stream’, in which cells are sensitive to information about the identity of objects, and an occipito-parietal ‘dorsal stream’, in which cells are sensitive to spatial information (Ungerleider & Mishkin 1982). From the traditional cognitive perspective, the ventral stream builds a representation of ‘what’ is in the environment, while the dorsal stream builds a representation of ‘where’ things are. However, the dorsal stream does not appear to contain any unified representation of the space around us, but rather diverges into a number of substreams each specialized towards the needs of different kinds of actions (Stein 1992; Andersen et al. 1997; Wise et al. 1997; Colby & Goldberg 1999; Matelli & Luppino 2001). For example, the lateral intraparietal (LIP) area is concerned with the control of gaze (Snyder et al. 1997), represents space in a body-centred reference frame (Snyder et al. 1998a, b) and is strongly interconnected with parts of the oculomotor system including the frontal eye fields (FEFs) and the superior colliculus (Paré & Wurtz 2001). In contrast, the medial intraparietal (MIP) area is involved in arm reaching actions (Ferraina & Bianchi 1994; Kalaska & Crammond 1995; Snyder et al. 1997), represents target locations with respect to the current hand location (Graziano et al. 2000; Buneo et al. 2002) and is interconnected with frontal regions involved in reaching, such as the dorsal premotor cortex (PMd; Johnson et al. 1996; Marconi et al. 2001).

These observations are consistent with the proposal that the major role of the dorsal visual stream is not to build a unified representation of the world, but rather to mediate various visually guided actions (Goodale & Milner 1992). It may therefore be part of the system for action specification (Fagg & Arbib 1998; Kalaska et al. 1998; Cisek & Turgeon 1999; Cisek 2001; Passingham & Toni 2001), processing visual information to specify potential actions of various kinds: LIP cells specify potential saccade targets; MIP cells specify possible directions for reaching, etc. Furthermore, the dorsal stream represents not only a single unique movement...
that has already been selected, but rather offers a variety of options to choose from multiple saccade targets (Platt & Glimcher 1997; Kusunoki et al. 2000) as well as multiple reaching movements (Cisek et al. 2004). It does not, of course, represent all possible movements. As one proceeds along the dorsal stream, one finds an increasing influence of attentional modulation, with information from particular regions of interest enhanced while information from other regions is suppressed (Desimone & Duncan 1995; Treue 2001). The result is that the parietal representation of external space becomes increasingly sparse as one moves away from striate cortex (Gottlieb et al. 1998). In other words, only the most promising targets for movements make it so far to be represented in the parietal cortex. From this perspective, the phenomenon of selective attention is seen as an early mechanism for action selection (Allport 1987; Neumann 1990; Tipper et al. 1992, 1998), reducing the volume of information that is transformed into action-related representations.

As mentioned, parietal cortical areas are strongly and reciprocally interconnected with frontal regions involved in movement control. LIP is interconnected with FEF, MIP with PMd and primary motor cortex (M1), AIP with ventral premotor cortex (PMv), etc. (Matelli & Luppino 2001). As a result, the frontal-parietal system may be viewed as a set of loops spanning over the central sulcus, each processing information related to a different aspect of movement (Pandya & Kuypers 1969; Jones et al. 1978; Marconi et al. 2001). If these regions are involved in representing potential actions, as assumed here, then they appear to do so in tandem. For example, potential reaching actions are represented together by both MIP and PMd (Cisek et al. 2004; Cisek & Kalaska 2005). It is proposed that the competition between potential actions plays out in large part within this reciprocally interconnected fronto-parietal system. Within each cortical area, cells with different movement preferences mutually inhibit each other, creating a competition between distinct potential actions. This competition is biased by excitatory input from a variety of sources, including both cortical and subcortical regions. The influence of all these biasing factors modulates the activity in frontal and parietal neurons, with information favouring a given action causing activity related to that action to increase, while information against an action causes it to decrease.

Indeed, neurophysiological evidence for the modulation of fronto-parietal activity by ‘decision factors’ is very strong. For example, recent studies on decision making show that LIP activity correlates not only with sensory and motor variables, but also with decision variables such as expected utility (Platt & Glimcher 1999), local income (Sugrue et al. 2004), hazard rate (Janssen & Shadlen 2005) and relative subjective desirability (Dorris & Glimcher 2004). More generally, variables traditionally considered as sensory, cognitive or motor appear to be mixed in the activity of individual cells in many regions, including prefrontal cortex (PFC; Hoshi et al. 2000; Constantinidis et al. 2001), premotor cortex (Romo et al. 2004; Cisek & Kalaska 2005), FEF (Thompson et al. 1996; Gold & Shadlen 2000; Coe et al. 2002), LIP (Platt & Glimcher 1997; Shadlen & Newsome 2001; Coe et al. 2002) and the superior colliculus (Basso & Wurtz 1998; Horwitz et al. 2004). Such mixing of variables is difficult to interpret from the perspective of distinctions between sensory, motor and cognitive systems, and it has led to persistent debates about the functional role of specific cortical regions. For example, some studies have shown that neurons in LIP area respond only to stimuli which capture attention, leading to its interpretation as a ‘salience map’ (Colby & Goldberg 1999; Kusunoki et al. 2000; Bisley & Goldberg 2003). However, other studies have shown that these activities are stronger when the stimulus serves as the target of a saccade (as opposed to a reach), leading to the interpretation of LIP as a representation of intended saccades (Snyder et al. 1997, 1998a,b, 2000). These competing interpretations have been the subject of a long and vibrant debate. However, from the perspective of the affordance competition hypothesis, both interpretations are correct: neural activity in fronto-parietal regions correlates with sensory and motor variables because it is involved in the specification of potential actions using sensory information, and it is modulated by decision variables (including salience/attention) because a competition between potential actions is influenced by various sources of biasing inputs.

There are many potential sources from which biasing inputs might originate. Since action selection is a fundamental problem faced by even the most primitive of vertebrates, it probably involves neural structures which developed very early and have been conserved in evolution. A promising candidate is the basal ganglia (Mink 1996; Kalivas & Nakamura 1999; Redgrave et al. 1999; Frank et al. 2007; Hazy et al. 2007), which are strongly interconnected with specific cortical areas (Alexander & Crutcher 1990a; Middleton & Strick 2000) and exhibit activity that is related to both movement parameters (Alexander & Crutcher 1990b,c) and decision variables such as reward (Schultz et al. 2000) and expectation (Lauwereyns et al. 2002). However, it is also probable that action selection involves brain structures which have become particularly developed in recent evolution, such as the PFC of primates. The PFC is strongly implicated in decision making (Bechara et al. 1998; Kim & Shadlen 1999; Fuster et al. 2000; Miller 2000; Rowe et al. 2000; Tanji & Hoshi 2001), which may be viewed as an aspect of advanced action selection. Neurons in the dorsolateral prefrontal cortex (DLPFC) are sensitive to various combinations of stimulus features, and this sensitivity is always related to the particular demands of the task at hand (di Pellegrino & Wise 1991; Hoshi et al. 1998; Rainer et al. 1998; Kim & Shadlen 1999; Quintana & Fuster 1999). Prefrontal decisions appear to evolve through the collection of ‘votes’ for categorically selecting one action over others, as demonstrated by studies of saccade target and reach target selection (Kim & Shadlen 1999; Tanji & Hoshi 2001). Of course, the PFC is not a homogeneous system but a diverse collection of specialized regions, including some which appear to be involved in the aspects of working memory (Fuster & Alexander 1971; Bechara et al. 1998; Petrides 2000; Rowe et al. 2000). Here, we include only a very simplified account of one particular subregion of PFC, the DLPFC.
What role might the ventral visual stream play within the functional architecture of figure 1? Cell responses in anterior inferotemporal (IT) cortex are sensitive to the features of a currently viewed stimulus (Desimone et al. 1984; Tanaka et al. 1991), and to the behavioural context in which this stimulus is presented (Eskandar et al. 1992). These results have been taken to implicate IT in object recognition. However, it may also serve a more humble role. Studies on animal behaviour over the last hundred years have shown that many kinds of behaviours are elicited by simple combinations of particular stimulus features, which ethologists referred to as ‘sign stimuli’ (Tinbergen 1950; Hinde 1966). Neural responses in IT cortex are compatible with a putative role in sign stimulus detection, which could serve as a front-end input to action selection via direct projections from temporal cortex to prefrontal regions (Saleem et al. 2000). Thus, an early role of what is now the ventral stream may have been the detection of the stimulus combinations that were relevant for selection of actions in a particular behavioural context, and this may have eventually evolved into the sophisticated object recognition ability of modern mammals.

3. A COMPUTATIONAL MODEL OF REACHING DECISIONS

The broad concepts outlined in §2 can be translated into more concrete and testable hypotheses through a mathematical model of the neural processes which may implement action specification and selection in the mammalian cerebral cortex. A model of the cortical mechanisms which specify reaching movements and select between them has been described by Cisek (2006) and is summarized briefly here (see also the electronic supplementary material).

Figure 2a illustrates the circuit model and suggests how its elements may correspond to specific cortical regions. Since the model focuses on visually guided reaching actions, it includes some of the main cortical regions involved in reaching behaviour, such as the posterior parietal cortex (PPC), dorsal premotor cortex (PMd), primary motor cortex (M1) and prefrontal cortex (PFC). These were chosen as a subset of the complete distributed circuits for reaching control, sufficient to demonstrate a few central concepts. Other relevant regions not currently modelled are the supplementary motor areas, somatosensory cortex and many subcortical structures including the basal ganglia, red nucleus, etc. The input to the model consists of visual information about target direction and a signal triggering movement onset (GO signal), and the output is the direction of movement. The control of the overt movement is not simulated here (for compatible models of execution, see Bullock & Grossberg 1988; Houk et al. 1993; Kettner et al. 1993; McIntyre & Bizzi 1993; Bullock et al. 1998; Cisek et al. 1998).

In the model, each neural population was implemented as a set of 90 mean-rate leaky-integrator neurons, each of which is broadly tuned to a particular direction of movement. All the weights are fixed to resemble the known anatomical connections between the modelled regions. Within each population, neurons with similar tuning excite each other, while neurons with dissimilar tuning inhibit each other. Between populations, neurons with similar tuning excite each other through reciprocal topological connections. Noise is added to all neural activities. For details of the model’s implementation, see the electronic supplementary material.

In the model, neural populations do not encode a unique value of a movement parameter (such as a single direction in space), but can represent an entire distribution of potential values of movement parameters (e.g. many possible directions represented simultaneously). This proposal is related to the attention model of Tipper et al. (2000), the ‘decision field’ theory of Erhagen & Schöner (2002) and the ‘Bayesian coding’ hypothesis (Dayan & Abbott 2001; Sanger 2003; Knill & Pouget 2004). It suggests that a given population of cells, each with a preferred value
of a particular movement parameter, one can interpret the activity across the population as something akin to a probability density function of potential values of that parameter. Sometimes, the population may encode a range of contiguous values defining a single action, and at other times, several distinct and mutually exclusive potential actions can be represented simultaneously as distinct peaks of activity in the population (figure 2b).

The strength of the activity associated with a particular value of the parameter reflects the likelihood that the final action will have that value and is influenced by a variety of factors including salience, expected reward, estimates of probability, etc. This hypothesis predicts that activity in the population is correlated with many decision variables, as observed in frontal (Kim & Shadlen 1999; Gold & Shadlen 2000; Hoshi 2004) and parietal cortices (Platt & Glimcher 1999; Shadlen & Newsome 2001; Coe et al. 2002; Glimcher 2003; Dorris & Glimcher 2004; Sugrue et al. 2004; Janssen & Shadlen 2005).

The model suggests that sensory information in the dorsal visual stream is used to specify the spatial parameters of several currently available potential actions in parallel. These potential actions are represented simultaneously in frontal and parietal cortical regions, appearing as distinct peaks of activity in the neural populations involved in sensorimotor processing (Platt & Glimcher 1997; Cisek et al. 2004; Cisek & Kalaska 2005; figure 2b). Whenever multiple peaks appear simultaneously within a single frontal or parietal cortical region, they compete against each other through mutual inhibition. This is related to the biased competition mechanism in theories of visual attention (Desimone 1998; Boynton 2005). To state it briefly, cells with similar parameter preferences excite each other, while cells with different preferences inhibit each other. This basic mechanism can explain a variety of neural phenomena such as the inverse relationship between the number of options and neural activity associated with each (Basso & Wurtz 1998; Cisek & Kalaska 2005), narrowing of tuning functions with multiple options (Cisek & Kalaska 2005) and relative coding of decision variables (Roesch & Olson 2004).

Since neural activities are noisy, competition between distinct peaks of activity cannot follow a simple ‘winner-take-all’ rule, or random fluctuations would determine the winner each time, rendering informed decision making impossible. To prevent this, small differences in the levels of activity should be ignored by the system. However, if activity associated with a given choice becomes sufficiently strong, then it should be allowed to suppress its opponents and conclusively win the competition. In other words, there should be a threshold of activity above which a particular peak is selected as the final response choice. This is consistent with sequential sampling models of decision making (Usher & McClelland 2001; Mazurek et al. 2003; Reddi et al. 2003; Smith & Ratcliff 2004; Bogacz et al. 2007) which propose that decisions are made when neural activity reaches some threshold. In the model, this threshold emerges from the nonlinear dynamics between competing populations of cells (Grossberg 1973; Cisek 2006; see electronic supplementary material).

Finally, the model suggests that the competition which occurs between potential actions represented in the fronto-parietal system is biased by a variety of influences from other regions, including the basal ganglia (Redgrave et al. 1999) and PFC (Miller 2000; Tanji & Hoshi 2001) which accumulate evidence for each particular choice (figure 1). Here, only the influence of PFC is modelled, although it is probable that basal ganglia projections play a significant role in action selection (Frank et al. 2007; Houk et al. 2007). Several studies have shown that some cells in lateral prefrontal cortex (PFC) are sensitive to conjunctions of relevant sensory and cognitive information (Rainer et al. 1998; White & Wise 1999; Miller 2000; Tanji & Hoshi 2001), and that they gradually accumulate evidence over time (Kim & Shadlen 1999). Many studies have suggested that orbitofrontal cortex and the basal ganglia provide signals which predict the reward associated with a given response (Schultz et al. 2000), which could also serve as input to bias the fronto-parietal competition.

The operation of the model can be most easily understood in the context of a particular task. For example, figure 3a shows a reach-decision task in which the correct target was indicated through a sequence of cues: during the spatial-cue (SC) period, two possible targets were presented, and during a subsequent colour-cue (CC) period, one of these was designated as the correct target. In the model, the appearance of the spatial cue causes activity in two groups of cells in PPC, each tuned to one of the targets. Mutual excitation between nearby cells creates distinct peaks of activity, which compete against each other through the inhibitory interactions between cells with different preferred directions. Owing to the topographic projections between PPC and PMd, each peak appears in PMd as well, although they are weaker in the lower PMd layers (compare layers PMd1 and PMd2). These two peaks continue to be active and compete against each other even after the targets vanish, owing to the positive feedback between layers. At the same time, activity accumulates in the PFC cells selective for the particular location–colour conjunctions. The colour cue is simulated as uniform excitation to all PFC cells preferring the given colour (in this case, PFC<sub>Y</sub>), and it pushes this group of PFC cells towards stronger activity while the other is suppressed. In the model, this is equivalent to a decision. Finally, once the GO signal is given, activity is allowed to flow from PMd<sub>Y</sub> into M1 and the peak of the M1 activity is taken to define the initial direction of the movement.

The simulation reproduces many features of neural activity recorded from the dorsal premotor and primary motor cortex of a monkey performing the same reach-decision task (Cisek & Kalaska 2005). As shown in figure 3a, PMd cells tuned to both spatial targets were active during the SC, and then during the CC, one of these became more strongly active (predicting the monkey’s choice), while the other was suppressed. Note how the activity was weaker while both options were present, consistent with the hypothesis that the two groups of cells exert an inhibitory influence on

Phil. Trans. R. Soc. B (2007)
each other. As in the model, these phenomena were seen more strongly in the rostral part of PMd than in the caudal part. The model also exhibits sustained activity (‘working memory’) because after the targets are removed (second black line in the simulation images), target information is maintained in both PPC and PMd (figure 3a, b).

Figure 3c shows a variation of the task in which the CC is presented before the SC. In this case, no directionally tuned activity appears in PMd during the CC period, and after the spatial targets are presented, there is sustained activity corresponding only to the correct target. Thus, the neural activity is determined not by the sensory properties of the stimulus (which are the same as in figure 3a), but by the movement information specified by the stimulus. However, note that immediately after the SC, there is a brief burst towards the incorrect target, in both the neurons in rostral PMd and in the PMd population in the model (figure 3c). One might be tempted to classify this as a pure ‘sensory’ response. However, at least in the model, this burst is more correctly described as a brief representation of a potential action, aborted quickly in light of the prior information provided by the colour cue. Again, this is seen most strongly in the rostral part of PMd, in both the data and the model.

In addition to reproducing qualitative features of neural activity during the reach-decision tasks of Cisek & Kalaska (2005), the model produces important psychophysical results on the spatial and temporal characteristics of human motor decisions. For example, it is well known that reaction times in choice tasks increase with the number of possible choices. This can be explained by the model (figure 4a) because the activity associated with each option is reduced as the number of options is increased (compare model PMd activity in figure 3a with figure 3b), and it therefore takes longer for the activity to reach the decision threshold. Furthermore, it has also been shown that reaction time is not only determined by the number of targets, but also by their spatial configuration. For example, Bock & Eversheim (2000) showed that
activity exceeded 0.75. Proposed by other models which involve a gradual
produced this (figure 4 times longer and more broadly distributed. The model
reproduces the important finding that reducing the
quality of evidence for a given choice makes reaction
time when three cues are presented 80
different magnitudes. The decision latency was calculated as
simulations (each with two targets) using a CC cue of
(2007)
Phil. Trans. R. Soc. B
This finding is difficult to account for with models in
which the options are represented by discrete groups of
angle, but shorter if two targets are closer together.
representations of potential actions.
cerebral cortex may implement a competition between
interaction with the world, which requires animals to
to select between them. This framework is used to interpret
neural data from the primate cerebral cortex,
suggesting explanations for a number of important
neurophysiological phenomena. A computational
model is presented to illustrate the basic ideas of the
hypothesis and suggest how neural populations in the
cerebral cortex may implement a competition between
representations of potential actions.

4. DISCUSSION
This paper describes a theoretical framework called the
‘affordance competition hypothesis’, which suggests that behaviour involves a constant competition
between currently available opportunities and demands
for action. It is based on the idea that the brain’s basic
functional architecture evolved to mediate real-time
interaction with the world, which requires animals to
continuously specify potential actions and to select
between them. This framework is used to interpret
neural data from the primate cerebral cortex,
suggesting explanations for a number of important
neurophysiological phenomena. A computational
model is presented to illustrate the basic ideas of the
hypothesis and suggest how neural populations in the
cerebral cortex may implement a competition between
representations of potential actions.

The mathematical model presented above shares a
number of features with existing models of decision
making. For example, it is similar to a class of models
called ‘sequential sampling models’ (Roe et al. 2001; Usher & McClelland 2001; Mazurek et al. 2003; Reddi et al. 2003; Smith & Ratcliff 2004), which propose that
decisions are made by accumulating information for a
given choice until it reaches some threshold. In some
models, the evidence is accumulated by a single process
(e.g. Smith & Ratcliff 2004), in some it is collected by
separate processes which independently race towards

Figure 4. Latency effects. (a) Simulated reaction time during
tasks with one, two, three or four targets presented for 1.3 s,
followed by a single correct target for 0.1 s, followed by the
GO signal. Reaction times were calculated as first time after
the GO signal that any neuron in the M1 population exceeded
an activity threshold of 1.5. The mean and standard error are
shown for N = 300 replications in each condition. (b) Simulated reaction time when cues are presented for
0.8 s followed by a single target for 0.3 s prior to the GO. The bars show mean ± s.e. of reaction time in four conditions:
when three cues are presented 80° apart, two cues 160° apart,
two cues 80° apart or no cue at all. N = 100 in each condition.
(c) Distributions of decision latency computed during
simulations (each with two targets) using a CC cue of
different magnitudes. The decision latency was calculated as
the time between the CC cue and the first time any PMd3 cell
activity exceeded 0.75. N = 200 for each condition.

reaction time in a reaching task is similar with two or
five targets as long as they subtend the same spatial
angle, but shorter if two targets are closer together.
This finding is difficult to account for with models in
which the options are represented by discrete groups of
neurons, but is easily reproduced in a model such as the
present one, in which movements are specified by a
continuous population (figure 4b). The model also
reproduces the important finding that reducing the
quality of evidence for a given choice makes reaction
times longer and more broadly distributed. The model
produces this (figure 4c) through the same mechanism
proposed by other models which involve a gradual
accumulation to threshold: that with weaker evidence,
the rate of accumulation is slower and the threshold is
reached later in time, and therefore variability in
accumulation rate produces broader distributions of
reaction times (Carpenter & Williams 1995; Ratcliff
et al. 2003; Smith & Ratcliff 2004).

The model also explains several observations on the
spatial features of movements made in the presence of
multiple choices. For example, Ghez et al. (1997)
showed that when subjects are forced to make choices
quickly, they move to targets randomly if they are
spaced farther than 60° apart (‘discrete mode’), and
in-between them if the targets are close together
(‘continuous mode’), as shown in figure 5a. The model reproduces all of these results (figure 5b).
When two targets are far apart, they create multiple
competing peaks of activity in the PMd–PPC popu-
lations, and the decision is determined by the peak that
happens to fluctuate higher when the signal to move is
given. However, if the targets are close together, then
their two corresponding peaks merge into one owing to
the positive feedback between cells with similar
parameter preferences (a similar explanation has been
proposed by Erlhagen & Schöner 2002). In a related
experiment, Favilla (1997) demonstrated that the
discrete and continuous modes can occur at the same
time when four targets are grouped into two pairs that
are far apart, but each of which consists of two targets
close together (figure 5c). This is also reproduced by
the model (figure 5d; except for an additional central
bias exhibited by human subjects). With four targets,
peaks corresponding to targets within each pair merge
together and then the two resulting peaks compete and
are selected discretely.
the threshold (Roe et al. 2001; Reddi et al. 2003), and in some the independent accumulators inhibit each other (Usher & McClelland 2001; Bogacz et al. 2007). Some models separate the decision process into serial stages (e.g. Mazurek et al. 2003) and in some it occurs when a single population exhibits a transition from biased competition to binary choice (Wang 2002; Machens et al. 2005). While the present model shares similarities with these, it extends their scope in an important way. In all of the models of decision making described above, the choices are predefined and represented by distinct populations, one per choice. In contrast, the present model suggests that the choices themselves emerge within a population of cells whose activity represents the probability density function of potential movements. In other words, the model describes the mechanism by which the choices are defined using spatial information. In this sense, it is related to the models of Tipper et al. (2000) and Erlhagen & Schöner (2002), which also discuss continuous specification of movement parameters within a distributed representation. To summarize, the present model may be seen as combining three lines of thought: (i) sequential sampling models of accumulation of evidence to a threshold, (ii) models of a phase transition from encoding options to binary choice behaviour (see electronic supplementary material), and (iii) models of action specification within a distributed population. It also suggests a plausible manner in which these concepts can be used to interpret neural data in specific cortical regions.

The model presented here makes a number of predictions which distinguish it from many other models of decision making. First, it focuses on decisions about actions (as opposed to sensory discrimination) and suggests that these are made within the very same neural circuits that control the execution of those actions. These circuits are distributed among a large set of brain regions. In the case of visually guided reaching, decisions are made within the fronto-parietal circuit that includes both PMd and parietal area MIP. In the specific mathematical formulation described above, the competition between actions uses information from PFC, but the decision first appears in PMd, in agreement with data (Wallis & Miller 2003). However, the broader framework of the affordance competition hypothesis does not impose any rigid temporal sequence in which decisions appear in the fronto-parietal system. Each population in the network is proposed to involve competitive interactions, and biasing influences can modulate this competition in different places. Since cortico–cortical connections are bidirectional, if a decision begins to emerge in one region, then it will propagate outward to other regions. For example, decisions based on sensory features such as stimulus salience may first appear in parietal cortex and then influence frontal activity. In contrast, decisions based on abstract rules may first be expressed in frontal regions and propagate backward to PPC. Thus, decisions are proposed to emerge as a ‘distributed consensus’, which is reached when a competition between representations of potential actions is
unbalanced by the accumulation of evidence in favour of a given choice.

Although the mathematical model presented here is similar in some ways to previous models of decision making, it is based on a somewhat unusual theoretical foundation. The affordance competition hypothesis, illustrated schematically in figure 1, differs in several important ways from the cognitive neuroscience frameworks within which models of decision making are usually developed. Importantly, it lacks the traditional emphasis on explicit representations which capture knowledge about the world. For example, the activity in the dorsal stream and the fronto-parietal system is not proposed to encode a representation of objects in space, or a representation of motor plans, or cognitive variables such as expected value. Instead, it implements a particular, functionally motivated mixture of all of these variables. From a traditional perspective, such activity appears surprising because it does not have any of the expected properties of a sensory, cognitive or motor representation. It does not capture knowledge about the world in the explicit descriptive sense expected from cognitive theories and has proven difficult to interpret from that perspective (see above). However, from the perspective of affordance competition, mixtures of sensory information with motor plans and cognitive biases make perfect sense. Their functional role is not to describe the world, but to mediate adaptive interaction with the world.

In summary, instead of viewing the functional architecture of behaviour as serial stages of representation, we view it as a set of competing sensorimotor loops. This is by no means a novel proposal. It is related to several theories which describe behaviour as a competition between actions (Kornblum et al. 1990; Hendriks-Jansen 1996; Toates 1998; Prescott et al. 1999; Ewert et al. 2001), and as discussed above, to a number of philosophical proposals made throughout the last hundred years. The present discussion is an attempt to unify these and related ideas with a growing body of neurophysiological data. It is suggested that a great deal of neural activity in the cerebral cortex can be interpreted from the perspective of a competition between potential movements more easily than in terms of traditional distinctions between perception, cognition and action (Cisek 2001). It is not suggested that distinctions between perceptual, cognitive and motor processes be discarded entirely (they are certainly appropriate for interpreting primary sensory and motor regions), but only that other conceptual distinctions may be better suited to understanding central regions.

Figure 6 provides a schematic of the conceptual differences between the affordance competition hypothesis and the traditional frameworks of cognitive neuroscience. Traditional frameworks tend to view brain function as consisting of three basic classes of neural processes (figure 6a): perceptual systems, which take sensory information and construct internal representations of the world (e.g. Marr 1982); cognitive systems, which use that representation along with memories of past experience to build knowledge, form judgments and make decisions about the world (Newell & Simon 1972; Johnson-Laird 1988; Shafir & Tversky 1995); and action systems, which implement the decisions through planning and execution of movements (Miller et al. 1960; Keele 1968). Each of these broad classes can be subdivided into subclasses.

Phil. Trans. R. Soc. B (2007)
For example, perception includes different modalities such as vision, which can be subdivided further into object recognition, spatial vision, etc. Likewise, cognition includes processes such as working memory storage and retrieval, decision making, etc. These conceptual classes and subclasses are used to define research specialities, categorize scientific journals and interpret the functional role of specific brain regions.

Here, a different taxonomy of concepts is proposed (figure 6b). Brain function is seen as fundamentally serving the needs of interactive behaviour, which involves two classes of processes: action specification, which use sensory information to define potential actions and guide their execution online; and action selection, which help to select which potential action will be performed at a given moment. Each of these can be subdivided further. For example, action specification can be divided into the specification of different kinds of actions, such as reaching, which involves spatial vision, inverse kinematics, etc. Action selection includes processes such as visual attention which selects information on the basis of sensory properties, as well as decision making which selects potential actions on the basis of more abstract rules. Note that many of the same concepts appear within both taxonomies, albeit in a different context. For example, vision of space is seen as closely related to object recognition in figure 6a, but in figure 6b, they are thought of as contributing to very different behavioural abilities.

It is proposed here that the taxonomy of figure 6b is better suited to interpret neural activity in many brain regions because it more closely reflects the basic organization of the nervous system. Several aspects of brain anatomy are reflected in figure 6b, such as the distinction between tectal and striatal circuits, dorsal and ventral visual streams and the divergence of parietal processing towards different kinds of actions (of course, the specification and selection systems are not completely separate: as described above, mechanisms for action selection must influence activity related to specification at many loci of sensorimotor processing throughout the dorsal stream). Furthermore, one may view the relationships between the conceptual classes and subclasses in figure 6b as reflecting, at least to some extent, the phylogenetic relationships between them. For example, one can speculate that processes such as ‘object recognition’ evolved as specializations of older mechanisms of decision making which did not explicitly represent the identity of objects but simply detected particular features, called ‘sign stimuli’ (Tinbergen 1950; Hinde 1966). A classification of concepts which aims to reflect, at least to some extent, the phylogenetic features, called ‘sign stimuli’ (Tinbergen 1950; Hinde 1966). A classification of concepts which aims to reflect, at least to some extent, the phylogenetic relationships between them.

The author wishes to thank Andrea Green and Steve Wise for their helpful comments on various versions of this manuscript. This work was supported by the New Emerging Teams grant NET-54000 from the Canadian Institutes of Health Research and a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES


MacLean, P. D. 1973 A triune concept of the brain and behaviour. Toronto, Canada: University of Toronto Press.


Mead, G. H. 1938 The philosophy of the act. Chicago, IL: The University of Chicago Press.


