Chapter V

General Discussion:

Relation of Eye Movements to Other Movements
Several lines of evidence indicate that the knowledge gained about control of saccadic eye movements can be generalized to movements of other limbs producing complex behaviour. For example, the influence of waiting period before a target on response time is the same for movements of the eyes (Findlay, 1981) as it is for movements of the limbs (Niemi & Näätänen, 1981); the influence of the number of movements that increases the latency of first movement in a sequence of saccades (Zingale & Kowler, 1987) follow the same trend observed for speech and typing (Sternberg, Monsell, Knoll & Wright, 1978). The slowing down of a response that has been observed following errors of vocal or manual movements (Rabbitt, 1966, Rabbitt & Phillips, 1967) has also been observed for saccadic eye movements (Cabel et al., 2000, Schall & Taylor, 1998), suggesting that the influence of the consequence of a movement on the latency of a subsequent movement seems to be independent of the effector. In this general discussion I will try to relate my findings to the general problem faced by brain in sequences of movements.

**Concepts and findings in other sequential behaviors**

Complex motor actions typically consist of simpler movements assembled in a proper sequence to achieve a goal; for example to drink water from a glass resting on a table, we first stretch the hand to reach the glass and then retract the hand to bring the glass to the mouth. Sequential movements are also evident in the temporally structured innate behaviour of animals (Colombo, Eickhoff & Gross, 1993, Kermadi, Jurquet, Arzi & Joseph, 1993); for example in different gait patterns in quadrupeds such as walking, trotting and galloping. In some cases the sequential organization of movements can be quite complex such as those that occur during grooming sequences (Aldridge, Berridge, Herman & Zimmer, 1993, Berridge & Whishaw, 1992), or in birdsong (Konishi, 1985). How this sequencing is accomplished in brain is a fundamental yet unresolved issue.

Pavlov and other early observers postulated that each element in a series of actions provides the excitation of the next using feedback mechanism. This mechanism has been called stimulus-response reflex chaining or simply response chaining (Bain, 1868, James, 1890) or associative chaining (Wickelgren, 1969). One of the assumptions of response chaining theory was that subsequent movements in a series were thought to
be triggered by feedback from the preceding movement. The response chaining theory was called into question by the observation that a movement sequence can be performed in the absence of sensory feedback. Severing the dorsal root of the spinal cord, through which afferent sensory signals pass, neither rendered the patient nor the experimental animal motionless; they continued producing movements in sequence (Bard, Fleury, Teasdale, Paillard & Nougier, 1995, Nougier, Bard, Fleury, Teasdale, Cole, Forget, Paillard & Lamarre, 1996, Sanes, 1990, Taub & Berman, 1968). In order to resolve this confound it was suggested that the problem of serial order could be understood better by observing why serial order collapses (Lashley, 1951). For this Lashley used data on sequencing errors, particularly speech error, in which early and later elements of a sequence mistakenly exchanged positions, to infer that neural representations for all elements of a planned sequence are simultaneously active before sequence production. Lashley believed that sensory input comes into play only after the plans for organized sequential movements are made. This has been reiterated by other investigators suggesting that the centrally programmed command sequences are “structured before the movement begins and allows... the entire sequence to be carried out uninfluenced by peripheral feedback” (Keele, 1968).

When the order of subsequent movements is known before the initiation of the first movement, serial motor behaviour can take place using preprogrammed commands without relying upon feedback control that compares the ongoing response and the goal using visual information of the moving limb or the internal representation of the sensory consequence of the movement. Therefore the latency of the first response is supposed to increase with the number of responses in a sequence since the preparation time of each response contributes to the time spent before the onset of the first response. Sternberg and his colleagues tested this hypothesis asking well-trained typists to produce a short sequence of keystrokes as rapidly as possible recalling from memory (Sternberg et al., 1978). They found (1) the latency, i.e. the time to first keystroke after the onset of the start signal increased as a linear function of sequence length; (2) mean inter response interval (IRI) increased as a function of sequence length; (3) the ratio of latency to mean IRI was greater than one; (4) for a given length of sequence, some IRIs were longer than others, but the longer IRIs occurred in different positions depending on sequence length;
and initiation of the final item was faster. A similar increase in reaction times of the first utterance with the length of the sequence to be produced was observed in speech tasks or spoken immediate serial recall (ISR) tasks that involved the presentation of a list of familiar items (e.g. digits, letters, words) that the participant was asked to recall in the correct order (e.g. Conrad, 1958, Cowan, 1994).

A rudimentary mechanism explaining the delay in response was first explained by Henry and Rogers (1960), who asked subjects to respond to the onset of a signal either by lifting the hand from a switch; or by lifting the hand from a switch and performing one more action; or by lifting the hand from a switch and performing two more actions. The movement latency, i.e. the time of initiation of movement after signal onset, increased as the number of motor task to be performed increased. They explained their finding using the metaphor of a memory-drum. The memory-drum model suggested that information about the forthcoming movement sequence was held in memory, and became accessible for production by a mechanism abstractly similar to the rotating drums that memory psychologists then used to automate presentation of stimuli at variable rates. They proposed that the net drum rotation rate slows as the sequence to be recalled becomes more complex. Subsequent studies (Anson, 1982, Christina, Fischman & Vercruyssen, 1982) have confirmed these finding; and the inference drawn by Henry and Rogers (1960) holds true for other tasks too involving speech (Sternberg, Wright, Knoll & Monsell, 1980), key pressing (Rosenbaum, Inhoff & Gordon, 1984), handwriting (Hulstijn & van Galen, 1983) and saccadic eye movements (Zingale & Kowler, 1985).

However, one prediction of the memory-drum model was that performance rate of all elements in a sequence is expected to be uniformly slow. This prediction has been falsified in multiple studies on human subjects (Heuer, 1988, MacKenzie & Van Erde, 1990, Verwey, 2003). In their choice reaction task (cRT) Rosenbaum, Saltzman & Kingman (1984) instructed subjects to minimize the time between appearance of the discriminative (choice) signal and the completion of one of two possible finger-tapping sequences. Participants were instructed to choose between “i” vs. “I”, “ir” vs. “IR”, or “irm” vs. “IRM”, where “i”, “r”, “m” denoted key presses of the index, ring and middle fingers of the left hand, respectively, and “I”, “R”, “M” denoted key presses of the index, ring and middle fingers of the right hand, respectively. Subjects learned to associate one
visual signal (O) with one sequence and another signal (X) with the other. The main results were that latency of first key press $T_1$ increased with the number key presses in the sequence; and that the mean time of second key press $T_2$ was longer when that response was embedded in a sequence of three than when embedded in a sequence of two. Rosenbaum, Inhoff & Gordon (1984) introduced contextual effects by manipulating the position of the sequence difference. In different blocks of trials, the choices were between “irm” vs. “Irm”, “irm” vs. “iRm”, or “irm” vs. “irM”. Therefore the task incorporated uncertainty in response at three different positions of the sequence separately. The hypothesis was that if subject had chosen between sequences from two independent hierarchy trees, then the latency $T_1$ of the first response in the common sequence (“irm”) should be independent of the position of the uncertain response in the sequence. But result showed that $T_1$ increased with the position of selection cue in the sequence. These results are in accord with the notion of hierarchical control suggesting that constituents are successively unpacked to their lowest levels and the higher the level of transition in hierarchy, the longer the time it takes for a transition.

**The Modeling of movement sequences**

Over half a century ago Karl Lashley (1951) called serial order in behaviour “…the most important and most neglected problem of cerebral physiology”. Lashley noted that because individual response can appear in a variety of contexts, any given response may be associated with more than one subsequent response. In such cases, information limited to the recent response provides an ambiguous cue for selection of the future response. He considered the response chaining account untenable and postulated instead that the production of serial behaviour involves that parallel activation of a set of responses, which together comprise some “chunk”, so that responses are internally activated before being externally generated. In addition to representations of individual responses, the actor must also have access to a broader representation of temporal context, a “schema” that somehow encodes the overall structure of the intended response sequences. A “schema for action” independently selects which response, of those activated, to produce at which time. In his word “…elements of the sequence
are...partially activated before the order is imposed upon them in expression suggest that some scanning must be at play in regulating their temporal sequence”.

Since Lashley a noteworthy progress in the field of organized serial motor behaviour of animal and human has been made. These studies have mainly addressed the issues how are acquired sequences learned, represented and executed using various kind of behavioral paradigm, for example vocalization, typing, manual pointing etc. The results of these studies have suggested that complex movement sequences are controlled hierarchically. Hierarchical control is often depicted as a tree-like branching structure, where the highest level of planning is best described by how an action achieves an objective; lower levels of hierarchy are dedicated to translating a goal into movements. In the context of serial behaviour, the highest level representation of a sequence to be produced corresponds to the sequence’s main constituent, lower level representations correspond to lower level constituents and so on (Fuster, 1989, Grafman, 1995, Miller, Galanter & Priham, 1960, Schank & Abelson, 1977).

Routine activities in everyday behaviors also incorporate serial order of actions that humans perform frequently without needing to pay attention to the task at hand, for example making the daily cup of breakfast coffee while still being half asleep or planning the day ahead. Figure 5.1 illustrates that actions seemingly grouped together into subroutines, which are again part of a larger routine. Thus we may have a processing system structured as a hierarchy of processing units. Rather than being organized by a unitary schema, behaviour here has been described as involving the coordination of multiple schemas, each associated with different levels of temporal structure.

Hierarchical processing structure of actions warrants a mechanism of selection and serial ordering (scheduling) of responses activated in parallel. A mechanism, called ‘contention scheduling’ for the resolution of competition among responses to generate serial ‘routine’ action was proposed (Norman & Shallice, 1986, Shallice, 1982). Shallice envisaged lateral inhibition among competing responses may be a reason that the most intended response (active schema) to be the one which gains control of effector systems. Norman & Shallice’s (1986) widely acknowledged dual-systems theory of the control of action claims that two distinct systems contribute to the expression of a complex sequential behaviour: One is an automatic conflict resolution system (Contention
Figure 5.1. Hierarchical representation of a routine sequential task. The processing system of making a cup of tea is structured as a hierarchy of processing nodes or units with each node at a lower level representing a simple action and nodes at higher level representing progressively complex actions. Behaviour here has been described as involving the coordination of multiple schemas associated with different levels of temporal structure.
Scheduling, or CS), which selects one action from amongst the myriad of actions possible at any moment in time. CS is argued to function autonomously during the control of routine behaviour, i.e., when performing highly overlearned tasks such as one’s daily breakfast routine as illustrated in Figure 5.1. However, in deliberate behaviors (less familiar circumstances, novel tasks or dangerous situations), a higher level executive system (the Supervisory System, or SS) may exert control through modulation of CS. A useful analogy with regard to the relationship between these two systems is that of a horse (CS), which is doing all the actual work of locomotion, and a rider (SS), who can decide upon the path to follow in unusual circumstances.

The Interactive Activation Network (IAN) (Cooper & Shallice, 2006) and Simple Recurrent Network architecture (SRN) (Botvinick & Plaut, 2004) are two influential computational models that have been proposed in recent time to explain our routine serial behaviour and its breakdown. While the IAN model employs “goal nodes” in a hierarchy that gate activation flow to the lower level until the preceding action has been completed, the SRN model casts doubt on the functional roles of both schemas and goals and the need for assuming hierarchical structure. SRN (Figure 5.2) is a network in which each output is fed back as one component of a high-dimensional input (or other pre-output) stage. After extensive sequence-specific learning, this feedback signal combines with other state information to create a distinctive context for eliciting the correct next output. Another recent empirical work (Ruh, Cooper & Mareschal, 2005) suggests that such an all-or-nothing approach to a specific task being either routine or not might be misleading. The results instead support the conclusion that the dual systems interact in an extremely flexible manner, with the contribution of the SS, at every point in time, being dependent on factors such as the local complexity of the task and the amount of experience with this and similar tasks.

In his original proposal of contention scheduling Shallice (1982) suggested that one and only one schema become dominant at any given time, but he did not mention how a schema, once dominant might become de-activated. Therefore there remained the risk of endless repetition of dominant response. Moreover, their model postulated direct connections from activated response schemata to the effector, leading to a requirement of one schema to completely suppress all others in the response competition to prevent any
Figure 5.2. The three layered architecture of the simple recurrent network (SRN) model comprising a set of processing units. The activation of each unit is based on excitation and inhibition received from units linked to it through weighted synapse-like connections. The first layer of the model carries a pattern of activation representing some input to the system. Here for example, the visual features of the ‘fixated object’ currently at the focus of visual attention or features of a ‘held object’ the actor currently has in hand. The fixated object is interpreted as the target of an action; the held object is interpreted as the implement to be used. Activation from the input layer propagates to an intermediate hidden layer, in which each hidden unit is connected to every other. The hidden layer transforms the input information, sending a pattern of activation to an output layer whose units together represents the system’s response to the input; here for example, a movement directed to the ‘fixated object’.

Figure 5.2
chance of interference. Estes (1972) postulated that after some period of activity the dominant schema enters into a refractory period, enabling the next schema in the sequence to become active. Rumelhart and Norman (1982) extended this concept to fit transposition errors in serial order typing behaviour into their model. Transposition errors occur when any two adjacent letters are switched. In the hierarchical R&N model (Figure 5.3.A) specific sequences are represented as “word schema”, which corresponds to a word (or part of words). When a word is to be produced, nodes representing the letters in the word are equally activated in parallel by word-to-letter connections. The “key-press schemata” receive information about the current finger position from the effector system. Each letter schema inhibits the activation of all letters schemata that follow it. Such lateral inhibition gradually reduces the excitation from word schema across letters schemata. Thus the parallel activation generates serial output by inducing an activation gradient over the letter nodes.

Parallel representation underlying much of our learned serial behaviour has been a basis for the proposal (Grossberg, 1978a, Grossberg, 1978b) of another class of parallel sequence production models that have since come to be known as competitive queuing (CQ) models (Bullock & Rhodes, 2003). Such models (Figure 5.3.B) follow two basic assumptions: (1) more than one plan representation can be simultaneously active in a planning layer; (2) the most active plan is chosen in another layer called the competitive choice layer. The planning layer contains nodes representing possible sequence elements, such as letters of the alphabet A to Z in case of typing behaviour. To plan a sequence, a desired subset of these nodes is activated in parallel and the relative degree of activation is used to specify the relative priority of performance. At the onset of the gating signal, the active representations begin to compete for output via the choice layer. If the competition is fair, the most active plan-layer node will win the competition and thereby generate a corresponding output from the choice-layer, which initiates the action. A second effect of this output, mediated by an inhibitory connection from each output node to its corresponding plan-layer node, is deletion of the activity at whatever plan-layer node has just won to ensure that after the execution of a sequence, the plan-layer is empty and ready for the preparation of further sequences. Some experimental support in favor of the CQ type architecture is provided by the study of Averbeck et al. (2002), who trained
Figure 5.3. (A) Schematic diagram of the Rumelhart and Norman (RN) typing model. A 'word' is produced by excitation (black arrow) of each constituent 'letter' node representing key-press schemata in parallel. An activation gradient across letter nodes is induced by inhibition (grey arrows) from previous letters following the order in which they appear in a word. Thus the first letter receives no inhibition and later letters receive progressively more. Letter nodes are selected according to their level of activation for serial output. (B) Schematic of the competitive queuing (CQ) model in which an activation gradient across responses are formed in the parallel planning layer without making an ordered line. Each response competes with others for selection by lateral inhibition in the next layer. Selection takes place by a winner-take-all mechanism. Once a response is generated, the corresponding activation is deleted from the planning layer to prevent repetition of action.
monkeys to copy geometrical shapes shown on the screen as a sequence of strokes. The activity in area 46 of prefrontal cortex represented all serial elements of a movement sequence in parallel before the action began. The stroke most actively represented was produced first and the corresponding neural representation was deleted first, and so on, until the final stroke was produced.

An equivalent parallel representation of locations of potential targets is the concept of a salience map. As alluded to in the introduction, the hypothetical salience map in cortical and subcortical oculomotor system is a spatial representation of the conspicuousness of each object in a scene competing for selection as a target of visuo-spatial attention (Fecteau & Munoz, 2006, Itti & Koch, 2001). The assignment of weights to objects activated simultaneously in the salience map corresponds to the activation gradient in the planning layer of the CQ model. Along with this analogy, the allocation of attention to the most distinct object also prevents that object being selected twice through ‘inhibition of return’, which is conceptually similar to the deactivation of representation of a response in the choice layer of the CQ model. Thus the salience map provides a conceptual framework that accounts for how the next object is selected.

Although the salience map provides an explanation of how a distinct object can capture attention, how such covert selection translates into a shift of gaze is not so clear. Despite the wealth of evidence supporting causal relationship between covert shift of attention and overt eye movement (Awh, Armstrong & Moore, 2006, Deubel & Schneider, 1996, Nobre et al., 2000, Rizzolatti, Riggio, Dascola & Umilta, 1987), the distinctness between visual selection and motor preparation becomes apparent in tasks in which these two processes diverge, for example, when movement is either withheld (delayed or memory guided saccade task) or made away from the visually selected target (antisaccade task). Recent physiological evidence suggests that distinctness of processes corresponding to visual selection and motor preparation can be observed even at the level of single cells in the oculomotor network. Schall and his colleagues (Sato, Murthy, Thompson & Schall, 2001, Sato & Schall, 2003) examined the activity of single unit within frontal eye field (FEF) while monkey made saccades either toward or away from the distinct target in a search array. Two sets of neurons were found: one responsible for orientation of attention that initially selected the location of the target, and then
subsequently selected the final end point of saccade; and another responsible for oculomotor programming that only selected the end point of saccade. Further evidence of a dissociation between the orientation of spatial attention and gaze came from studies that showed how covert orientation of attention amplified the visual signal of the target in the absence of overt eye movement and movement planning in the frontal eye fields (Thompson, Biscoe & Sato, 2005) and the superior colliculus (Ignashchenkova, Dicke, Haarmeier & Thier, 2003).

The distinctness of visual and motor processes is compatible and even predicted by the salience model. This is because salience, which is determined by the distinctiveness of objects, is fundamentally driven by the visual context of the scene. Therefore, the saliency map hypothesis is essentially determined by bottom-up sensory mechanisms. Although the incorporation of feature biases (as proposed by Wolfe in the guided search model) allows the salience map to be modulated by top down control, the model does not incorporate the relevance of selected targets to the observer. Therefore a visual salience map may not suffice to explain sequences of saccades (or sequences of movements in general) that may take place in the real world in which the relation between objects and not their salience per se may determine their importance. For example, a nail in a cluttered visual scene might not be a salient object but in the context of a previous movement to lift a hammer might be accorded with a high priority that facilitates a subsequent movement towards it. Therefore I suggest that necessity of a combined representation of distinctiveness and relevance may form a priority (motor salience) map from which locations are selected as goals for saccades and or other types of movements generating goal directed movement sequences.

Although a speculation at this stage, I suggest that the motor related activity in oculomotor areas might be considered a critical node of such a priority map (at least for oculomotor sequences). Certain features of movement related activity makes this statement an attractive proposition. 1) The close relation between visually related cells in FEF allows them to receive information of salient objects in the environment. For example a movement to pick up a hammer cannot occur without the relevant object in the environment. 2) A well-documented relation between activity in these cells and motor output has been shown, as embodied in the LATER model. 3) In my thesis I suggest that
a LATER can be suitably modified to incorporate the programming of sequential movements by introducing two neural integrators that individually accumulate information at a steady rate to reach a common threshold. To instantiate control I assumed that when the activities of two neural integrators overlap in time they reduce the activity of each other by means of lateral inhibition. This inhibition introduces competition between the two integrators whose strength depends on the ratio of the activity of the two neural integrators. Thus the kernel emulates a competitive queuing model that maintains adequate temporal separation between two movements. Although not shown here, the kernel may be easily modified to specify directly the type of sequence to be produced in addition to their timing. In this respect, the kernel might be able to function in a manner analogous to the contention scheduling mechanism in the Shallice model. 4) Because the kernel parameters can be modified by cognitive context, the modeled activity of such cells can directly embody changes in priority, as demanded by error and subsequent correction. 5) The outcome or consequences of movements should be in a position to modify subsequent selection/planning of actions. Such a possibility was discussed in reference to the observed saccade-related bottleneck described in the last chapter.

In closing I hope that a combination of behavioral, modeling and electrophysiological studies may be able to test some of the ideas that have been put forth in my thesis in the hope that it will lead to a better understanding of how goal directed sequential movements may be programmed and controlled by our brains.