From Parallel Sequence Representations to Calligraphic Control: A Conspiracy of Neural Circuits

Daniel Bullock

Calligraphic writing presents many challenges for motor control, including: learning and recall of stroke sequences; critical timing of stroke onsets and durations; fine control of grip and contact forces; and letterform invariance under size scaling, which entails fine control of stroke directions and amplitudes during recruitment and derecruitment of musculoskeletal degrees of freedom. Experimental and computational studies in behavioral neuroscience have progressed toward explaining the learning, planning, and control exercised in tasks that share features with calligraphic writing and drawing. This article highlights component operations ranging from parallel sequence representations to fine force control. Treated in succession are: competitive queuing models of sequence representation, performance, learning, and recall; letter-size scaling and motor equivalence; cursive handwriting models in which sensory-motor transformations are performed by circuits that learn inverse differential kinematic mappings; and fine-grained control of timing and transient forces by circuit models that learn to solve inverse dynamics problems.

Key Words: handwriting, sequence representation, competitive queuing, neural network, cerebellum, adaptive resonance, vector integration, adaptive timing, chunking, inverse kinematics

Writing is a modality of linguistic communication and artistic expression—and a window on the neural substrates for representing and fluently executing serial plans. From the performance perspective, calligraphic writing reflects an ability to use a pen or brush to produce, in a highly repeatable manner, visible forms that approximate “ideal” letters (or ideograms) that are acknowledged as standards of beauty. Calligraphers can achieve letterform invariance—conserved relative proportions and slants—across large size variations and despite significant changes in “media” (which skeletal joints are used, what hand grip, what paper or other surface, what ink or paint, what stiff pen or compliant brush, etc.). Conserving letterform despite such variations presents a stringent challenge to the movement...
control system, and the requisite skill could take years to develop. No comprehensive model of the process exists, but several partial models provide clues to the ensemble of processes that are engaged during learning and performance of calligraphic writing. This review treats component process models formulated as neural networks to highlight abilities of distinct brain circuits.

**Calligraphic Instruction:**

**Prescribed Stroke Sequences and Recognition of Good Form**

Western calligraphic instruction usually begins by teaching students prescribed stroke sequences for production of full upper case and lower case alphabets for one historical style, e.g., “Irish uncial” or “Imperial Roman of the Trajan arch” (whereon letters were painted by calligraphers prior to the stonecutting; see Catich, 1968). After months of practice with one style at various sizes, further styles, requiring distinctive patterns of pen or brush use, are introduced, to sustain motivation, broaden and refine control and refine visual form discrimination. Stroke sequence instruction is often complete (no components omitted) and medium-grained (some specification of stroke overlaps and timing). The fine-grained elements of control, however, are hard to instruct, and so left to goal-directed self-discovery. Thus, a key complement to stroke sequence instruction is perceptual goal instruction. The novice calligrapher is given feedback that draws attention to small deviations between produced letterforms and the ideals. Thus the calligrapher is early taught to be a good critic, and then left alone to vary stroke vectors (amplitude and direction) and timing patterns during sequence production to produce letterforms that fall progressively closer to the ideals. This paper is restricted to the representation and skilled performance of prescribed stroke sequences.

**Sequence Representation in Competitive Queuing Models**

Behaviorist theories postulated that behavioral sequences are produced by sequential chaining, in which associative links allowed early elements of a sequence to elicit later elements. Thus the representation included a set of learned associative links, and the performance mechanism was iterated recall of the next element by the prior one. This idea was criticized and improved on by proponents of recurrent network models of sequence representation (e.g., Elman, 1995; Dominey, 1998; Beiser & Houk, 1998). In such models, a series of central states that are sequence-specific must be learned to mediate a sequence of recalls of subsequent elements in the sequence. This allows more than one sequence to be learned over the same alphabet of elements, but learning is slow, and internal recall of a sequence remains an iterative sequential operation.

In contrast, competitive queuing (CQ) models allow arbitrary sequences to be rapidly learned over the same alphabet of elements, and the internal recall of a sequence representation is a parallel operation. From the time of Lashley (1951), behavioral evidence has accumulated (Rhodes et al., 2004) to support the idea that parallel representation of elements constituting a sequence underlies much of our learned serial behavior. From speech and typing errors, Lashley inferred that there must be an active “co-temporal” representation of the items constituting a forthcoming sequence. He also inferred that item-item associative links might be
unnecessary in, and even a hindrance to, the learning of many sequences defined over a small finite alphabet. He did not, however, address questions of mechanism, such as: What is the nature of the parallel representation? How is the relative priority of simultaneously active item representations “tagged”? What limitations are inherent in this representation of order? What mechanisms convert the parallel representation to serial action? All four questions were addressed, without any reliance on item-item associations, in Grossberg’s (1978a, b) proposal of a class of parallel sequence representation and performance networks that have come to be known as CQ models (Houghton, 1990; Bullock & Rhodes, 2003). These neural network models postulate that a standing parallel representation, of all the items constituting a planned sequence, exists prior to initiation of performance of the first item. As explained in Figure 1, this parallel representation works in tandem with an iterated choice process to generate a sequential performance.

To date, CQ-compatible neural models have been successfully applied in many domains of learned serial behavior, including: eye movements (Grossberg & Kuperstein, 1986); recall of short novel lists (Boardman & Bullock, 1991; Page & Norris, 1998) and highly practiced lists (Rhodes & Bullock, 2002b); cursive handwriting (Bullock et al., 1993b); working memory storage of sequential inputs (Bradski et al., 1994); word recognition and production (Grossberg, 1986; Hartley & Houghton, 1996; Gupta & MacWhinney, 1997); language production (Dell et al., 1997; Ward, 1994); and music learning and performance (Mannes, 1994; Jacobs & Bullock, 1998; Page, 1999). Despite such successful behavioral applications, the status of CQ as a *neurobiological* model has been uncertain because of a lack of directly pertinent observations. Recent neural recordings in frontal cortical and other planning areas, however, have strikingly supported four key predictions of CQ models as originally proposed in Grossberg (1978a, b). Notably, the primate electrophysiological study of Averbeck et al. (2002; 2003) showed (1) that prior to initiating a serial act (of using a cursor to draw a geometric form with a prescribed stroke sequence), there exists in prefrontal area 46 an active parallel (simultaneous) representation of each of the strokes planned as components of the forthcoming sequence. Small pools of active neurons code each stroke. Also, (2) the relative strength of activation of a stroke representation (neural pool) predicts its order of production, with higher activation level indicating earlier production. As the sequence is being produced, (3) the initially simultaneous representations are serially deactivated in the order that the corresponding strokes are produced. In the figure drawing task of Averbeck et al. (2002), each neural pool’s deactivation started half way through the interval during which the corresponding stroke was drawn. Several studies (Averbeck et al., 2002; Basso & Wurtz, 1998; Cisek & Kalaska, 2002; Pellizzer & Hedges, 2003) of neural planning sites also show (4) partial activity normalization: the amount of activation that is spread among the plans grows more slowly than the number of plans (in the sequence), and eventually stops growing.

Normalization, which is essential for effective use of relative activation level to code relative item order, is readily achieved by competitive interactions among simultaneously active plans (Grossberg, 1978a, b). The resulting low upper-bound on the number (e.g., five) of plans that can be simultaneously active in a motor working memory for sequences is an established property of human working
Figure 1—A competitive queuing (CQ) network and associated cellular dynamics.

(A): All CQ models have at least two layers, a parallel planning layer and a competitive choice layer. The parallel planning layer contains nodes representing possible sequence elements, such as letters of the alphabet A through Z. To prepare a planned sequence, a desired subset of these nodes is activated in parallel (such as nodes representing the letters that spell the Australian greeting “GIDAY”) and the relative degree of activation is used to control the relative priority of performance. At the onset of a gating signal, the active representations begin to compete for output via the choice layer. If the competition is fair, then the most active plan layer node will always win the competition, and thereby generate a corresponding output from the choice layer, which initiates the chosen action. A second effect of this output, mediated by a recurrent inhibitory pathway from each output node to its corresponding plan layer node, is deletion of activity at whatever plan layer node has just won. For a two-item sequence, iteration of this choose–perform–delete cycle assures that an element’s initial relative activation level in the planning layer implicitly codes its relative priority in the forthcoming sequence, and that after the second choice, the plan layer will be empty, and thus ready for preparation of further sequences. For use with sequences longer than two, the planning layer must be designed so that deletion of any node’s activity leaves invariant the rank ordering of the remaining node activations. If nothing interrupts the feedback and iterated choice processes, then production of a planned sequence can be very fluid. Note that CQ needs no item-to-item associative links.

(B): A simulation of cellular dynamics in the plan layer of a normalized CQ model (Boardman & Bullock, 1991) during production of a five-item sequence, such as “GIDAY.” These simulation traces correspond remarkably well with cell firing recorded a decade later by Averbeck et al. (2002) in area 46. Each simulation trace marks the activation history of one of the sequence element representations during the interval from just before element 1’s performance to just after production of the entire sequence. Note that prior to movement there are five positive traces corresponding to five simultaneously active representations. Adapted with permission from Rhodes & Bullock (2002b).
memory (cf. Cowan, 2001; Miller, 1956). Besides explaining working memory limitations, competitive normalization can explain an empirical generalization known as the sequence length effect on latency (SLEL): for novel or lightly practiced sequences performed from memory under reaction-time conditions, the time to initiate performance of a prepared sequence increases as a function of the number of items in that sequence (Sternberg et al., 1978; Boardman & Bullock, 1991; Klapp, 1996; Rhodes & Bullock, 2002b). In normalized CQ models, the SLEL occurs because more items imply less activation for each item, including the first item, and because the level of activation during the preparatory interval is a strong determinant of latency. The lower the activation, the further the representation from the response-production threshold, and the longer the time needed to exceed threshold following a “go” stimulus.

**Long-Term Sequence Memory and Rapid Recall During Preparation**

Two advantages of the parallel representation of sequences used in CQ models become clear when we consider the need to form separate long-term memories for each of many sequences defined over a finite alphabet of elements, e.g., the 26-letter Roman alphabet. Across the set of memories, an item, e.g., “O,” will be followed by many, perhaps all, of the remaining 25 elements of the alphabet. In non-CQ models that use associative, item-item links as the basis for forming sequence memories, it becomes a very difficult problem for any given sequence to ensure that linking occurs from “O” to the correct one of the 25 potential successors, especially when some successors are much more probable than others in the set of learned sequences. Because CQ models do not use item-item associative links as a primary basis for long-term memory (LTM) of sequences, CQ models completely avoid this problem. A second major advantage is that the CQ parallel representation format—essentially a standing pattern of analog values—is the same format used by the brain to represent image information, e.g., the kind of information used to perceive and recognize faces, objects, and scenes. Therefore, any bases that the brain possesses for forming LTMs of image information are, in principle, applicable to the CQ format for parallel representation of sequences. Thus, theories of LTM for CQ models are often based on the idea that distinct neural network nodes, conceptualized as small ensembles of cells, can be recruited by a learning process to serve as both recognizers and recallers of distinct parallel analog patterns. In such models, control circuitry is needed to ensure that small but critical differences in the patterns—e.g., reordering of the relative activation levels of two item representations—lead to recruitment of different pattern recognizer nodes. Because the class of competitive learning networks known as adaptive resonance theory (ART) networks (e.g., Grossberg, 1980; Carpenter, 2001) is distinguished by such control circuitry, variants of ART networks often form the LTM components of CQ models (cf. Grossberg, 1986; Houghton, 1990; Mannes, 1994; Hartley & Houghton, 1996; Gupta & MacWhinney, 1997; Page, 1999; Rhodes & Bullock, 2002b).

Typically, ART-like networks are interpreted as models of cortical and thalamo-cortical circuits, because such circuits include the reciprocal pathways required to implement the bi-directional (bottom-up and top-down) signaling that enables ART-like networks to build up large stable databases of often similar (i.e.,
overlapping) yet distinguishable patterns defined over an alphabet of items (or image features). A key question is whether *sub-cortical* brain structures, which do not have ART-like architectures, might also provide substrates for LTMs of sequences. Two candidate structures are the basal ganglia and the cerebellum, both of which are known to be involved in procedural learning. Recent data, reviewed in Hikosaka et al. (2000), have strongly implicated the basal ganglia in initial sequence discovery and assembly, but less so in sequence recall from LTM; in contrast, the lateral cerebellum appears unnecessary for initial sequence learning, but has been strongly implicated in sequence recall at advanced stages of practice (Lu et al., 1998). These complementary results suggest that at least one non-ART-like neural architecture—the cerebellum—is important for sequence LTMs. This conclusion coheres with anatomical evidence (Dum & Strick, 2003; Kelly & Strick, 2003) that the cerebellum, long wrongly considered as solely a low-level motor coordinator, includes a segregated lateral circuit that receives inputs from, and sends outputs to, area 46 of the prefrontal cortex—the same area in which Averbeck et al. (2002) found evidence for a parallel sequence representation strikingly similar to that predicted by CQ models.

What is the role of the cerebellar input to prefrontal area 46? One hypothesis compatible with CQ models and the data of Averbeck et al. (2002) was mathematically formalized and implemented on a computer as the N-STREAMS model (Rhodes & Bullock, 2002b), which is schematized in Figure 2. The name “N-STREAMS” is an acronym for Neural Substrates That Rehearse, Encode, And Memorize Sequences. In this adaptive neural circuit model, perceived sequential inputs are represented (cf. Bradski et al., 1994) in a declarative working memory, WMd, that uses the standard CQ format: neural representation of the first input item gets highest activation, the second input’s representation gets next highest activation, etc. On decision to imitate the represented sequence (cf. Sekuler et al., 2003), a voluntary transfer occurs from this declarative WM to a plan WM, WMp (cf. Rowe et al. 2000). This *chosen plan* WMp forms the plan layer of a CQ network (Figure 1) that sits at the core of the N-STREAMS model (Figure 2), and so it uses a parallel representational format similar to the declarative WMd. In the theory and the computer model, the voluntary transfer between the two WMs serves four functions: (a) it transforms a declarative representation into a plan representation; (b) following a CQ-generated performance, a new transfer enables the WMp (which has been emptied during performance, see Figure 1) to be re-initialized by the still-loaded WMd, with each new transfer affording another rehearsal by the CQ system of whatever sequence is stored in WMd, including any novel sequence not yet stored in LTM; (c) each WM transfer generates signals that facilitate chunk learning (one form of LTM) by an ART-like cortical circuit; and (d) each WM transfer also generates signals that guide learning (a second form of LTM) within the non-ART-like cerebellar circuit. These two types of LTM included in N-STREAMS are complementary.
Figure 2—Macrocircircuit illustrating the global architecture of the N-STREAMS model. Only major components and links between them are depicted here, but the associated simulation model shows how these major parts can be fully implemented as neural networks. The N-STREAMS model captures the hypothesis that several substrates capable of sequence representation combine their outputs to compose the stream of behavioral outputs. At the core of N-STREAMS is a competitive queuing (CQ) system (see Figure 1). This includes a plan working memory (WMp) and a choice field capable of choosing (for performance) the most active remaining plan in WMp. The CQ core is augmented by a declarative WM (WMd) and a closely associated adaptive cortical chunking system, which learns to recognize and recall the parallel sequence representations that pass through the WMs. Sites of learning between WMd and the cortical chunking subsystem are shown as semicircles. Additional sites of learning are located in the other major augmentation of the CQ core, namely the cerebellar side-loop. The text emphasizes the upper cerebellar side-loop, which interacts with the CQ plan layer (WMp). The lower cerebellar side-loop is not treated in the main text. It illustrates the cerebellum’s output-to-input recurrence (see also Figure 4), which allows the N-STREAMS model to utilize item-to-item associative links as a subsidiary sequencing mechanism. Output-to-input recurrence also provides the cerebellum with a basis for implicit forward modeling, by allowing the next cerebellar output to be conditionalized, in effect, on the expected results of the current cerebellar output. Put another way, output-to-input recurrence spares the system from the errors it would make if the cerebellum failed to take account of the expected consequences of its current outputs while it is generating its next outputs. Adapted with permission from Rhodes & Bullock (2002b).
Cortical chunk learning recruits a new node to serve as a compressed LTM representation, and recognizer, of each novel sequence registered in the WMd. Because of reciprocal adaptive links between the chunk representation stage and the WMd, the system can be used for both recognition and recall of sequences, and for the kind of representational error-checking that makes ART-like systems able to learn stable representations of many similar patterns defined over an alphabet of items or features. After sequence chunks become established by rehearsal, they can serve as distinctive “cognitive context” representations for cerebellar learning and recall. The model’s cerebellar module learns to respond to chunk activation (from any input source) by effecting a direct, automatic, rapid, and chunk-specific parallel loading of the appropriate CQ-type sequence representation into the planning WMp. Such rapid parallel loading from procedural LTM in the cerebellum anticipates (i.e., precedes in time) the results of the slower cortical recall operation, which can also load the WMp (but only via the WMd and explicit voluntary transfer). Once learning enables the cerebellar output signals that mediate parallel loading to approach their asymptotic strengths, the rapid loading operation causes the CQ subsystem to pre-select the first element of the sequence in preparation for performance. Simulations (Rhodes & Bullock, 2002b) of these learning-dependent interactions illustrate how the SLEL can be present early in learning a sequence yet disappear with moderate or extensive practice of that sequence, a phenomenon that has now been replicated in several typing-like sequence production tasks (Klapp, 1996; Verwey, 1996; review in Rhodes & Bullock, 2002b). A similar phenomenon could occur in handwriting (Teulings, 1996). Thus, the model formalizes the hypothesis that learned outputs of lateral cerebellum override a complexity–latency tradeoff that is otherwise characteristic of CQ, and thereby enables the prefrontal cortex (PFC) to speed sequence initiation even as it off-loads WMd.

This hypothesis raises an interesting conceptual issue. Why does the cerebellar output not simply bypass PFC (and WMp)? What function is served by a transcerebellar loop with origin in and return to PFC? Why not use the cerebellum only at a lower level, to learn a “motor tape” whose expression doesn’t require limited-capacity WMp? Such a bypass would allow the PFC to disengage as mediator of performance at high levels of practice. In fMRI studies, such PFC disengagement seems to occur during learning of simpler tasks, but disengagement did not occur in the figure drawing task of Averbeck et al. (2002). Why not bypass WMp in all cases? A key reason is that delivering the sequence representation to “working memory” allows the sequence plan to be reworked, i.e., allows its expression in action to be adjusted for myriad purposes, such as speed scaling of performance, selective verbal emphasis, replacement of one word by a more apt one, expressivity in song—or size scaling in handwriting.

An open question is the range of tasks that the N-STREAMS model can explain. In tasks where the subject is aware of the sequence and chunking is known to occur, a cerebellar contribution to LTM for learned sequences has been demonstrated (e.g., Lu et al., 1998). One recent neuroimaging report, however, has challenged the idea that the cerebellum is also involved in the implicit learning that occurs in the serial reaction time (SRT) task, in which subjects often show evidence of sequence learning despite having no awareness that a nonrandom
sequence has been presented and learned. In particular, Seidler et al. (2002) showed that throughout an initial learning phase that included a distractor task performed in parallel with the SRT task, there was no evidence of learning in performance, and also no evidence of cerebellar activations. On removal of the distractor task, performance suddenly improved, and cerebellar activations suddenly became detectable. They interpreted this correlation to mean that removal of the distractor task suddenly enabled previously masked cortical learning to gain expression via the cerebellum and thereby immediately enhance performance. Correlation, however, does not imply causation. The N-STREMS theory is fully consistent with the Seidler et al. (2002) data, but it offers a strikingly different interpretation: removal of the distractor task allows cortical chunk learning to finally gain access to the WMP, via which such learning immediately begins to have an impact on performance (even without a cerebellar assist to performance). Simultaneously, the transfers from WMd to WMP generate teaching signals for cerebellar learning, and it is these teaching signals that cause the cerebellar activations observed in the neuroimaging study. This prediction remains to be tested.

Size Scaling and Learned Inverse Kinematics for Tool Use

Skilled calligraphic writing exhibits approximate letterform invariance under 100x size scaling. Such size scaling implies the existence of plan representations that can efficiently recruit different joint actions depending on the desired scale (Lacquaniti et al., 1987) and effectors (Bernstein, 1967; Wright, 1990; Rijntjes et al., 1999). Issues of rate and size scaling have been addressed in a series of modeling studies based on the general class of voluntary kinematic production models called vector integration to endpoint (VITE) models (Bullock & Grossberg, 1988a). In VITE-type models, a central pattern generator subject to voluntary size and rate modulation continuously computes and integrates movement command vectors to produce graceful multi-joint movements characterized by bell-shaped velocity profiles and other basic kinematic signatures of human movement (Bullock & Grossberg, 1988a, b; Nagasaki, 1989; Zhang & Chaffin, 1999). These models explicate computational roles of diverse cortical and subcortical cell types during reaching (Bullock & Grossberg, 1991; Bullock et al., 1998; Cisek et al., 1998; Cisek, 2001), while also simulating both the discharge patterns of these neuron types and an increasing range of behavioral data, including cursive writing and viapoint movements (Bullock et al., 1993b; Bullock et al., 1999), combined discrete and rhythmic movements (Jacobs & Bullock, 1998; Sterndad et al., 2000), interceptive reaching (Dessing et al., 2002) and reach-grasp coordination (Ulloa & Bullock, 2003).

Data from the recent Averbeck et al. (2002) study, besides confirming CQ predictions noted above, verified two assumptions of the 1993 VITEWRITE model of cursive handwriting production (Bullock et al., 1993b; see also Schomaker et al., 1989; Schomaker & Van Galen, 1996): (1) a small number of straight-line stroke representations suffice as basis elements of the plan for production of curved graphical forms; and (2) for production of graphical forms with significant curvature, readout of the next stroke representation (from a sequence plan WM) should begin about halfway through the current stroke, i.e., near its peak velocity. The CQ-compatible VITEWRITE circuit (see Figure 3) explains how voluntary
Figure 3—An integrated model for scaling size and rate of cursive writing while preserving approximate letterform invariance. The macrocircuit shown combines elements from a VITEWRITE network for generating pen tip trajectories in spatial coordinates with an adaptive DIRECT network that transforms between spatial and motor coordinates by recruiting joint rotations adequate to produce desired spatial velocities. The learned inverse mapping (“inverse internal model”), from spatial to motor coordinates, is a posture-dependent differential mapping: it transforms spatial velocities (not positions) into joint rotation velocities (not postures). The learned forward mapping (“forward internal model”) transforms from postures to associated finger or pen tip positions. Although the name DIRECT alludes to the key inverse differential mapping needed for robust motor equivalence (see text), all DIRECT-type models also incorporate the internal forward mapping. The latter allows continuous prediction of end-effector position during intervals when visual feedback is blocked, or when movement is so fast that (slow) visual feedback is not helpful. The GRO signal is an internal size scaling signal, whereas the GO signal is an internal movement gating and rate scaling signal. Adapted with permission from Bullock et al. (1993a, b) and Fiala (1994).
size and speed scaling of cursive handwriting can occur *simultaneously* during performance while letterforms are preserved.

Although preservation under scaling occurs in the 1993 model, size scaling was limited by the modest range afforded by the three degrees of freedom included in that model. Of course, real size scaling *is* limited by such factors, and every calligrapher becomes adept at positioning him or herself vis-à-vis the writing surface so that writing the next letter will not require nearing a limit of the range of motion of a required degree of freedom. Still, a more general model should encompass not just handwriting but writing achieved with any set of skeletomotor degrees of freedom, including proximal arm joints (Lacquaniti et al. 1987), or even “foot writing” (e.g., Rijntjes et al. 1999). Figure 3 shows one way such a model might be constructed. The VITEWRITE circuit would operate in spatial (rather than motor) coordinates and effect behavior via a neural network capable of transforming VITEWRITE’s spatial pen-velocity commands into joint rotation-velocity commands. The requisite coordinate transformation could be learned and performed by self-organizing neural networks of the type formalized as Directions to Rotations Effector Control Transform (DIRECT) models (Bullock et al., 1993a; Fiala, 1994; Guenther & Barreca, 1997; Barreca & Guenther, 2001) and DIVA models for speech production (Guenther et al., 1998). In such networks, learning creates inverse differential kinematic mappings that can recruit, as the movement unfolds, novel combinations of skeletomotor joint rotations adequate to produce desired motions of a hand or a tool held in the hand. Such models enable real-time “motor-equivalence,” i.e., equivalent results from variable motoric means—a critical requirement for both tool use and size scaling. As such, the DIRECT class of vector integration to endpoint models also provides a basis for understanding how brain circuits can achieve the same performance characteristics as non-neural, yet behaviorally explanatory, models, such as the recent “optimal feedback” model (Todorov & Jordan, 2002; see also Latash et al., 2002). In particular, DIRECT’s motor equivalence capability enables on-line adjustments to perturbations of multi-joint movements that require high accuracy of tool positioning. The same features enable a hybrid VITEWRITE-DIRECT model to explain the main aspects of letterform invariance exemplified by calligraphic writing with brush or pen at the small scale of handwriting or the large scale of arm-writing.

The sensory-motor transform by which DIRECT models achieve motor equivalence makes distinctive predictions regarding the nature of *direction tuning* in motor cortex. In particular, the DIRECT model in Bullock et al. (1993a) predicted that the recruitment of cells in the motor cortex would prove to be a function of both the desired finger or pen direction and the current arm configuration—and not solely a function of the desired pen direction, as was assumed in many motor cortical data analyses prior to 1995. Thus, DIRECT predicted that motor cortical direction-tuning will be posture-dependent: a cell’s “preferred direction” will strongly depend on arm configuration, which necessarily changes across the writing space if the torso remains stationary. This key prediction was verified in a series of recent experimental (e.g., Scott & Kalaska, 1995; 1997) and computational modeling studies (Ajemian et al., 2000; 2001a; 2001b) of motor cortical coding of posture-dependent kinematics- and force-controlling signals. The integration proposed in
Figure 3 is compatible with these recent motor cortical analyses, as well as with most aspects of the extended VITEWRITE model of Contreras-Vidal et al. (1998) and recent reinforcement-learning and volitional gating models, e.g., Brown et al. (1999; 2004). The latter three models integrate cortical with basal ganglia circuits. This is highly pertinent, because the basal ganglia have been more strongly implicated than any other brain site in the gating of plan expression and in normal voluntary speed scaling that leaves direction and amplitude (spatial movement vectors) intact (Horak & Anderson, 1984; Hikosaka et al. 2000; Skinner and Garcia-Rill, 1990; Turner et al., 1998). Basal ganglia disease has also been implicated in distinct breakdowns of size and speed, e.g., in the diminutive handwriting (“micrographia”) seen in Parkinson’s disease (cf. Contreras-Vidal et al., 1998).

**Inverse Dynamics, Adaptive Timing, and Rapid Sequential Changes of Force**

Contreras-Vidal et al. (1997) showed how to combine a VITE model with an adaptive cerebellar model and a spinal circuit model to solve several key aspects of the inverse dynamics problem for a two-joint limb using both mono- and bi-articular muscles. This model was unique among otherwise similar models because of the realism and relative completeness of the spinal circuit model, which was an updated version of the factorization of length and tension (FLETE) model (Bullock & Grossberg, 1989; 1991). In particular, cerebellar learning was shown to be effective in a model that incorporated stretch, tendon, and joint receptors, mutually inhibitory Ia interneurons, Renshaw cell inhibition of motoneurons and Ia interneurons, and the size principle of motoneuron recruitment (see also Bullock, 2003 and van Heijst et al., 1998). Understanding the interactions among these ubiquitous mammalian circuit elements is critical for explaining graded force control over a full range from very small to very large forces. For example, the model formalized the hypothesis that the cerebellum learns to generate output signals that help the spinal cord effect a switch between a “postural” mode and a “fast movement” mode. Consistent with electrophysiological data (e.g., Henatsch et al., 1986), the model cerebellum’s phasic outputs perform a dual action in the spinal cord, by simultaneously exciting alpha motoneurons and inhibiting associated Renshaw cells that would otherwise recurrently inhibit the excited motoneurons. Thus, for the short duration of the phasic cerebellar output, the alpha-motoneuron pool and the stretch reflex operate at a higher “gain” than during times when Renshaw cells are not inhibited by phasic cerebellar output. This allows the system to benefit from the many postural virtues of Renshaw inhibition (cf. Bullock & Grossberg, 1989; Bullock & Contreras-Vidal, 1993) without suffering the problems that sustained Renshaw inhibition of motoneurons could pose for fast voluntary movements. This speed-enabling role of the cerebellum, which is supported in neuroimaging studies (e.g., Winsten et al., 1997) is abstractly akin to, but quite distinct from, the latency-shortening role proposed for the cerebellar projection to prefrontal cortex in the N-STREAMS model (Rhodes & Bullock, 2002b). In each case, the cerebellum allows a target circuit to overcome a rate-limiting property that would otherwise inhere in its operation.
Not treated in the Contreras et al. (1997) model was adaptive optimization of the timing of cerebellar outputs. Compatible models have been formulated, however, to explain the bases of cerebellar adaptive timing. Several of these models fall into the class of spectral timing models of the cerebellum (Bullock et al. 1994), and the version in Fiala et al. (1996) took the critical step of showing how intracellular calcium dynamics (triggered by glutamate acting via metabotropic receptors) can contribute directly to macroscopic behavioral timing. The most recent applications of such models (Ulloa et al., 2003; Bullock et al., 2003; Rhodes & Bullock, 2002a, b) explain how cerebellar adaptation of the timing and magnitude of control signals can help the motor cortex generate the minimal forces needed to prevent slip of an object from the hand during lifting with precision grip (Flanagan & Wing, 1997), whether the object’s surface is of high or low friction. Such cerebellar discovery of the minimal forces needed to avoid slip of a held object is pertinent to calligraphic writing, because a common problem with novice calligraphers is excess grip force.

![Figure 4](image-url)

Figure 4—Lattice-like structure of the cerebellar cortex. (A) shows the flow of signals from mossy fibers (MFs) through granule cells to their bifurcating axons, the parallel fibers (PFs), all within the cortex of the cerebellum. (B) depicts the convergence of PFs onto the giant planar dendrites of Purkinje cells. Purkinje cells inhibit deep cerebellar nucleus cells, which excite various action control centers. (C) schematizes the basic cerebellar circuit, including the pons, source of MFs, and the inferior olive, source of the climbing fibers whose discharges gate cerebellar learning. Also shown are Golgi cells, which are excited by MFs and PFs before feeding back inhibition to the granule cells. There are about as many granule cells in the cerebellum as there are cells in the remainder of the brain. Thus, there are many more granule cells than cells giving rise to MFs. Marr (1969) and Albus (1971) independently proposed that the granule-Golgi network in the cerebellar cortex performs a sparse combinatoric re-presentation of the active MF input vector. As a result, two similar MF activation vectors will give rise to highly dissimilar PF activation vectors. Such a transform would allow the cerebellum to learn highly nonlinear input–output functions. That is, highly similar inputs can learn to control highly dissimilar outputs. The proposed transform does not “invent arbitrary codes.” Rather, it reveals combinatoric information implicit in the input, consistent with the root meaning of “re-presentation.” Figure adapted with permission from Rhodes & Bullock (2002a).
For the cerebellum to generate the highly context-specific outputs (e.g., outputs suited to momentary arm postures as well as to frictional properties of both the pen and the paper) needed for advanced specialized skills, it must have access to a huge number of sui generis context representations. The major tenet of the classical Marr-Albus theory of information processing in the cerebellar cortex (Marr, 1969; Albus, 1971; Schweighofer et al., 2001) is that cerebellar processing creates such representations by expansively recoding the vectors of input signals carried by the mossy fibers that synapse on both granule and Golgi cells, as schematized in Figure 4. This “sparse context recoding” property is fully consistent with Fiala-type timing models (Fiala et al., 1996), but it is difficult to achieve in other models (e.g., Medina et al., 2000), which propose that the cerebellum’s granule and Golgi cell populations are dedicated to adaptive timing. At least as modeled in Medina et al. (2000), the latter proposal is inconsistent with five recent data sets (Shinkman et al. 1996; Raymond & Lisberger, 1998; Svensson & Ivarsson, 1999; Attwell et al., 2002; Kotani et al., 2003), whereas the same data sets confirm predictions of the Fiala-type model of cerebellar adaptive timing. This raises the intriguing possibility, formalized by Fiala et al. (1996), that intracellular calcium dynamics in at least one class of central neuron (Purkinje cells) has temporal properties that are matched to macroscopic behavioral requirements, which themselves depend on another calcium dynamic—the one at the heart of active muscle contraction.

This picture of cerebellar cooperation with kinematic generators as modeled with VITE or DIRECT provides a basis for understanding a broad range of behavioral and neurobiological data. Therefore, two recent papers that might be seen as challenges to the picture sketched above deserve mention. First, Karniel & Mussa-Ivaldi (2003) have suggested that the “system responsible for adaptation of movements to external forces might be unable to employ temporal representations (p. 10)” and instead always employs state representations (e.g., evolving information about joint positions and velocities). The theory sketched above suggests that the cerebellum does afford a “temporal representation.” Because it is nested within a combinatoric state representation, however, it does not satisfy the highly restrictive Karniel & Mussa-Ivaldi definition of a “temporal representation” as “some mechanism isolated from outside perturbations and the state of movement” (p. 19, emphasis added). Because of the cerebellum’s nesting of timing within state, the explicitly arranged decorrelation of state and timed force perturbations used in the Karniel & Mussa-Ivaldi study makes it impossible for the cerebellum to learn a timed sequence of compensatory responses. Thus, the cerebellar theory described here is consistent with the observed failure of human subjects to adapt. A second potential challenge is Spencer et al. (2003), which showed that cerebellar lesions disrupted timing of discontinuous but not continuous movements. The cerebellar state-dependent timing mechanisms described above would seem to be equally relevant for both types of movements, and indeed, there is strong evidence of normal cerebellar involvement in the success of continuous movements such as locomotion. One interpretation of Spencer et al. consistent with such evidence is that the only basis for adaptive timing in discrete movements is state-dependent cerebellar output, whereas for continuous movements there is another basis for timing that can compensate for loss of cerebellar assistance. Indeed, the very
presence of a reversal of direction in continuous movements allows an actor to use state-dependent—(e.g., time-to-contact dependent, and thus noncerebellar thalamus (Sun & Frost, 1998) dependent—goal switches in the kinematic generator (e.g., Jacobs & Bullock, 1998; Bullock, et al. 1999) as a basis for controlling the timing of continuous movements. This option is not available in discontinuous movements, which, by definition have a “dead time” between successive goal activations. To control their timing precisely, the cerebellum appears to be required.

Conclusions

Construction of a comprehensive, neurobiologically accurate, circuit model of an advanced example of fine sequential control of tool use, such as calligraphic writing with pen or brush, remains a task for the future. Mutually compatible neural circuit models of many of the constituent processes, however, are now in hand. A conspiracy (“breathing together”) of these models can explain emergent qualitative features of writing acts and the dynamics of more than 50 distinct physiologically identified cell types that exist in the planning and motor parts of the brain and spinal cord. Moreover, the models’ circuitry is consistent with neuroanatomical data from pathway-tracing studies. The models incorporate several types of on-line learning, mediated by distinct synaptic plasticity mechanisms that are known to be exhibited by the procedural learning system. These models highlight differentiated aspects of “learning by doing,” while explaining our ability to self-organize new skills that show incremental progress across temporally separated episodes of practice that can be freely interleaved with practice of other skills. In a few years, we have a realistic hope of simulating, with high neurobiological verisimilitude, how a clumsy pen wielder diligently transforms him or herself into an accomplished calligrapher.

Acknowledgments

I thank Prof. Robert Palladino of Reed College for instructing me (in 1973-1974) in techniques for producing a variety of styles of Western calligraphy. I also thank the organizing committee of the 2003 International Graphonomics Society meeting for inviting me to give a plenary talk to their congress. Preparation of that talk and this article was partially supported by National Institutes of Health grant R01 DC02852.

References


