David Marr: A Theory of the Cerebellar Cortex

A Model in Brain Theory for the "Galilean Combination of Simplification, Unification and Mathematization"

A.J. PELLIONISZ[1]

The brilliant but tragically short-lived – indeed, meteoric – career of David Marr was catapulted into orbit by his "Theory of the Cerebellar Cortex" (1969). One need not belabor the details of this well-publicized theory here (cf. Eccles 1973, Llinàs 1981, Ito 1984, Pellionisz 1985). Rather, three general observations on the theory will be briefly discussed.

First, the idea behind the theory was very uncomplicated. The chief neurons of the cerebellar cortex, the Purkinje cells, would "learn" to recognize a pattern over many of their input fibers (the parallel fibers) *if* such pattern would coincide with the deep depolarization of the Purkinje cell, evoked by the dense arbor of the climbing fiber which innervates it.

Second, a theory of the cerebellar *cortex,* by definition, focuses rather narrowly on a specific part of the cerebellar system, and is further focused on only one (hypothetical) facet of cerebellar function; plasticity. While a sharply focused view is characterized by dramatic highlights, it is probably impossible to understand the full role of the cerebellum in motor control without looking at it from a less spectacular but both broader and deeper perspective.

For example, from the viewpoint of the structure, cerebellar function may remain an enigma if the role of other cerebellar structures than only its cortex, e.g., the cerebellar *nuclei,* is not understood or specific cerebellar functions are not interpreted in the context of the structure of a whole sensorimotor system. Remember the old, hitherto unanswered question: "Can we make a real systems approach to cerebellar function without modeling the whole motor system?" (Arbib et al. 1968).

From a functional viewpoint, it is essential to point out that the cerebellum has been known for more than a century as the organ in the brain which acts as a motor coordinator (Flourens 1842); it is classic knowledge that its absence produces spatial and temporal dysmetria (Holmes 1939). True, the coordinator function of the CNS, just like any other CNS function, is endowed with a good degree of adaptability (cf. Llinàs and Pellionisz 1985). However, Marr's theory only provided a hypothesis for plasticity, and none for coordination –

[1] Department of Physiology and Biophysics, New York University, 550 First Avenue, New York, NY 10016, USA

Brain Theory
Edited by G. Palm and A. Aertsen

in effect assuming that motor learning was *the* function of the cerebellum. It was probably this limited character of the theory of cerebellar cortex that led Marr, shortly before his untimely death, to repudiate his theory as a utilizable explanation of what role the cerebellum plays in the motor system: "In my own case, the cerebellar study had two effects. On the one hand, it suggested that one could eventually hope to understand cortical structure in functional terms, and this was exciting. But at the same time the study has disappointed me, because even if the theory was correct, it did not much enlighten one about the motor system – it did not, for example, tell one how to go about programming a mechanical arm" (Marr 1982, p. 15).

Third, the idea has completed its course during the past two decades. Since the original concept of the theory by Brindley (1964), a rather large initial theoretical followership of the idea emerged; see Smolyaninov (1966), Eccles et al. (1967), Szentàgothai (1968), Grossberg (1969), Marr (1969), Albus (1971). Later, championed by the eminent experimentalist (see Eccles 1977), a campaign was launched to "prove" the theory. The early results, however, showed that "unfortunately, experimental testing of this hypothesis has failed to discover any significant modification even after some hundreds of parallel fibre-climbing fibre inputs to Purkinje cells with a 10-ms time discrimination (J.C. Eccles, D. Marr, N.H. Sabah, R.F. Schmidt, and H. Tabořikova, unpublished observations)" (Eccles 1973). Despite such discouraging results, merely because of the appeal of the theory, the campaign was enlisted by workers too numerous to name here. One and a half decades of intensive and devoted research later, two pieces of circumstantial "evidence" have emerged. After applying repeated parallel fiber – climbing fiber stimuli, one study reports a "depression" at the Purkinje cell level (Ito et al. 1982), while the other, based on a somewhat different experimental paradigm, shows an "enhancement" (Bloedel et al. 1983). These apparently inconsistent phenomena, neither of which lasts for more than a matter of hours, appear to be meager support for a theory which is based on the notion that the cerebellar cortex "learns" movements by permanently imprinting patterns into a set of connectivities of Purkinje neurons. At present, only one remaining major school of experimentation retains the "cerebellar learning paradigm" as its theoretical foundation.

Most certainly, it is admirably easy to put limiting qualifications on any major school of thought *from the retrospect of two decades*. Let us see, however, whether it is equally obvious to discern those qualities of Marr's theory of the cerebellar cortex, which created the spectacular success that his theory truly amounted to. For a definitive answer, it is tempting to accept the very conductive notion by Churchland (1986) – that the most critical test of a theory is whether it provides a "Galilean combination: the right sort of simplification, unification, and above all, a *mathematization*." Indeed, Marr's theory on the cerebellar cortex amply met each of these criteria.

In simplification, Marr's cerebellar theory even surpassed the Galilean viewpoint (*mutatis mutandis*). Identifying the function of the cerebellum as a uniquely simple operation – learning parallel fiber patterns by the Purkinje cells – was not unlike naming a planet around which the planetary system rotates. However, while the extraordinary simplicity of an assumption has its virtues, i.e. instant acceptance and universal acclaim, history documents that such axioms, even when shown inappropriate, are extremely difficult to dislodge. To shift the center of attention from a planet (or from plasticity) to a Sun (or to coordination) may take an inordinate amount of both time and exasperation.

Unification was perhaps the strongest virtue of Marr's theory of the cerebellar cortex. As he reminisced, his theory "suggested that one could eventually hope to understand cortical structure in functional terms, and this was exciting" (Marr 1982). An understatement. Indeed, by showing how the structure explains the function of *one third of the brain,* it was almost an entry into the promised realm of the inner sanctum of brain sciences. This was a refreshing contrast to the state-of-the-art of neuronal modeling, in which skilled applied mathematicians accounted for phenomena on rather minute fragments of nervous subsystems. His attempt at a grand unification of cerebellar structure and function was almost unprecedented, since such synthesis had been tackled only once before, by the pioneering notion that parallel fiber excitation volleys, arriving at Purkinje cells at different distances, would make the cerebellar cortex serve "as a clock in the millisecond range" (Braitenberg 1967). That earliest unification of micromorphological and overall functional features (beyond the problem of the demonstrably too short span of such timing), however, was based only on fairly trivial mathematization, and lacked some further unifying powers that Marr so deftly wielded in his theory. Two of the most important mergers that he accomplished were as follows: (1) Marr unified the knowledge of an anatomically distinct neuronal *circuitry* with a hypothesis at a *synaptic* level (that of Hebb 1949, who postulated that alteration of synaptic efficacies might subserve learning paradigms). (2) Equally important, Marr's theory of the cerebellar cortex unified the interpretation of the function of a *neuronal system* with theories of *abstract automata,* in the sense that through his theory the cerebellum appeared as a neuronal embodiment of Rosenblatt's (1959) learning machine, the Perceptron. Thus, from a wider perspective, Marr's theory – especially as exploited by Albus (1971, 1981) – was an attempt at unifying the fields of *Cybernetics* (cf. Wiener 1948) with biological studies of the cerebellum, which is at the forefront of *Neuroscience* (cf. Eccles et al. 1967).

Lastly, the *mathematization* of Brindley's (1964) germinal descriptive notion by Marr's quantitative elaboration must be addressed. As one of the most successful modelist in biology of all times comments: "It is not that most neurobiologists do not have some general concept of what is going on. The trouble is that the concept is not precisely formulated" (Crick 1979). As for Brindley's idea; for lack of mathematization it was dormant for years before Marr (1969), Grossberg (1969) and Albus (1971) matched the *concept* with a suitable mathematical *formalism.* This quantum jump elevated the notion into the league of theories, since a measure of the maturity and rigor of a field of science is the extent to which mathematics is utilized. By the "theory of cerebellar cortex" we have been given an example how brain research must progress towards exact synthesis through the sort of mathematization of biological concepts as achieved, for example, by Helmholtz (1896), Hodgkin and Huxley (1952) or Marr (1969).

Considering the mix of extraordinary merit and ordinary fallibility, what remains for us from Marr's theory of the cerebellar cortex?

It all depends on who we are. Naturally, in the field of experimental investigations, work always continues to "prove" or "disprove" *all* of the available theories, like people continue to savour all wine, old or new. Basing an experiment on a vintage theory is still better than working with no theoretical basis at all. Experimentalists can select from whatever theory is available (for recent reviews, see Eccles 1977, Llinàs and Simpson 1981, Ito 1984, Pellionisz 1985 or Pellionisz and Llinàs 1985).

For brain theory, however, the issue is not whether to prove or disprove theories with facts, if the old saying (attributed to Einstein) that "facts do not kill theories" is true. More advanced theories, like superior products of evolution, automatically *take precedence* over less potent earlier models; they all linger along, their fate being determined only by the rule of "the survival of the fittest".

Finally, for brain theorists, Marr's model of the cerebellar cortex remains forever a model for simplicity, unification, and mathematization; a high standard that one strives to equal, and perhaps to surpass.

References

Albus J (1971) A theory of cerebellar function. Math Biosci 10:25–61

Albus JR (1981) Brains, behavior and robotics. McGraw-Hill, New York

Arbib MA, Franklin GF, Nilsson N (1968) Some ideas on information processing in the cerebellum. In: Caianiello ER (ed) Neuronal networks. Springer, Berlin Heidelberg New York, pp 43–58

Bloedel JR, Ebner TJ, Qi-Xiang Yu (1983) Increased responsiveness of Purkinje cells associated with climbing fiber inputs to neighboring neurons. J Neurophysiol 50:220–239

Braitenberg V (1967) Is the cerebellar cortex a biological clock in the millisecond range? In: Fox CA, Snider RS (eds) Progress in brain research, vol 25. The cerebellum. Elsevier, Amsterdam, pp 334–346

Brindley GS (1964) The use made by the cerebellum of the information that it receives from sense organs. IBRO Bull 3:80

Churchland PS (1986) Neurophilosophy: Towards a unified understanding of the mind-brain. Bradford Books/MIT Press, Boston, Massachusetts

Crick FHC (1979) Thinking about the brain. Sci Am 241:219–232

Eccles JC (1973) The cerebellum as a computer: Patterns in space and time. J Physiol London 229:1–32

Eccles JC (1977) An instruction-selection theory of learning in the cerebellar cortex. Brain Res 127:327–352

Eccles JC, Ito M, Szentàgothai J (1967) The cerebellum as a neuronal machine. Springer, Berlin Heidelberg New York

Flourens P (1842) Recherches experimentales sur les proprietes et les fonctions du systeme nerveux dans les animaux vertebres, 2nd edn. Bailliere

Grossberg S (1969) On learning of spatiotemporal patterns by networks with ordered sensory and motor components. 1. Excitatory components of the cerebellum. Stud Appl Math 48:105–132

Hebb DO (1949) The organization of behaviour. Wiley, New York

Helmholtz H (1896) Handbuch der Physiologischen Optik. Voss, Leipzig

Hodgkin AL, Huxley AR (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol (London) 1176:500–544

Holmes G (1939) The cerebellum in man. Brain 63:1

Ito M (1984) The cerebellum and neural control. Raven Press, New York

Ito M, Sakurai MK, Tongroach P (1982) Climbing fibre induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar Purkinje cells. J Physiol (London) 324:113–134

Llinàs R (1981) Cerebellar modeling. Nature (London) 291:279–280

Llinàs R, Pellionisz A (1985) Cerebellar function and the adaptive feature of the central nervous system. In: Berthoz A, Melvill-Jones G (eds) Reviews of oculomotor research, vol I. Adaptive mechanisms in gaze control. Elsevier, Amsterdam, pp 223–231

Llinàs R, Simpson JI (1981) Cerebellar control of movement. In: Towe AL, Luschei ES (eds) Handbook of behavioral neurobiology, vol V. Motor coordination. Plenum Press, New York, pp 231–302

Marr D (1969) A theory of the cerebellar cortex. J Physiol (London) 202:437–470

Marr D (1982) Vision. A computational investigation into the human representation and processing of visual information. Freeman, San Francisco

Pellionisz A (1985) Tensorial brain theory in cerebellar modeling. In: Bloedel et al. (eds) Cerebellar functions. Springer, Berlin Heidelberg New York, pp 201–229

Pellionisz A, Llinàs R (1985) Tensor network theory of the metaorganization of functional geometries in the CNS. Neuroscience 16:245–273

Rosenblatt F (1959) Two theorems of statistical separability in the perceptron. Proceedings of Symposium on the mechanization of thought process. HMSO, London, pp 421–456

Smolyaninov VA (1966) Some special features of organization of the cerebellar cortex. In: Gelfand IM, Gurfinkel VS, Fomin SV, Tsetlin ML (eds) Models of the structural-functional organization of certain biological systems. MIT Press, Cambridge, pp 251–325 (Translated from Russian 1971)

Szentàgothai J (1968) Structuro-functional considerations of the cerebellar neuron network. Proc IEEE 56:960–968

Wiener N (1948) Cybernetics, or control and communication in the animal and the machine. MIT Press, Cambridge