Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
BioSystems 87 (2007) 82–95
BioSystems
www.elsevier.com/locate/biosystems

Perspectives on the Neuroscience of Cognition and Consciousness

Gerhard Werner *

Department of Biomedical Engineering, University of Texas, Engineering Science Building, Room 610, 1 University Station, C0800, Austin, TX 78712-0238, United States

Received 7 February 2006; received in revised form 23 March 2006; accepted 23 March 2006

Abstract

The origin and current use of the concepts of computation, representation and information in Neuroscience are examined and conceptual flaws are identified which vitiate their usefulness for addressing the problem of the neural basis of Cognition and Consciousness. In contrast, a convergence of views is presented to support the characterization of the Nervous System as a complex dynamical system operating in a metastable regime, and capable of evolving to configurations and transitions in phase space with potential relevance for Cognition and Consciousness.

Keywords: Cognition; Consciousness; Metastability; Phase space; Coordination dynamics; Computation; Representation; Information

1. Introduction

The conceptual framework of this essay is predicated on Searle's (1983, 1992) philosophical stance of biological naturalism: specifically, conscious states are entirely caused by lower level neurobiological processes; and they are "realized in the brain as features of the brain system, at a higher level than that of neurons and synapses" (Searle, 2004). The features designate subjective states of sentience or awareness. My main objective in the following is to propose a neurobiological process for the realization of the features: the issue at stake is to assemble neurobiological evidence that would support an account for the origin of the features' qualitative novelty as a state space transition, according to principles of nonlinear dynamics in complex physical systems.

In the context of this essay, I will consider Cognition and Consciousness as subjective states of sentience and awareness (Searle, 1992) and, when speaking of Neuroscience, I will limit the consideration of empirical data to the kinds obtained with neurophysiological and neuroanatomical methods, omitting for the purposes of this essay the enormously important areas of neurochemistry, molecular biology and neuro-genetics.

Before turning to the main objective, it will be necessary to clear the way and remove obstacles with foundational issues in Neuroscience which, I submit, lie in the way of meaningfully addressing issues in Cognition and Consciousness. This will complement Searle's (1998) list of philosophical obstacles to studying Consciousness scientifically. I take my cues from Hacking's (1985) "Styles of Scientific Reasoning", tracing currently prevalent discourse practices of Neuroscience from their enticing origins and initial plausibility to becoming sources of conceptual confusion and to precluding viable alternatives.

* Correspondence to: 4723 Cat Mountain Drive, Austin, TX 78731, United States. Tel.: +1 512 471 3604; fax: +1 512 471 0616.
E-mail address: gwer1@mail.utexas.edu.

0303-2647/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.biosystems.2006.03.007

2. With a nod to Foucault: a kind of Archeology of Neuroscience

The time frame for close scrutiny is quite short: with the outline of the classical Neuron doctrine firmly established, much basic Neuroanatomy and clinical Neuropsychology ascertained, and the essence of neurohumoral transmission and the ionic mechanism of neuronal electrical activity essentially in place, the field was by the mid 1940s ripe for assimilating novel ideas and technical advances that promised entirely new horizons for conceptualizing the Nervous System. These were forthcoming in rapid succession: Norbert Wiener's Cybernetics with the notion of feedback and control; Shannon's Information Theory; the invention of the Turing machine and the formulation of the Church-Turing's Thesis; the construction of the first large electronic computers, and von Neumann's invention of their programmability. Ling and Gerard demonstrated the use of micro electrodes to record electrical activity from individual neurons in the brain, and Adrian (1928) had already shown that single action potential spikes in peripheral nerve fibers transmit action potentials elicited by stimulation of their receptors.

Just imagine the goldmines that suddenly seemed to fall into the hands of system-oriented Neurophysiologists: single neurons and nerve axons delivering a binary code, seemingly just ready made for computing and information transmission in circuits and neural nets with feedback. For me who has experienced these years: the excitement was immense, and the promise seemed unlimited!

For exploring the scientific and social implications of these innovations, the Macy Foundation sponsored a series of annual conferences of the "Cybernetics Group", beginning 1943 and extending for the next 10 years; a virtually unparalleled undertaking. The attendance at the annual meetings varied from time to time, with W. McCulloch usually being in a leading role. Participants, usually 20 or so in number, were drawn from the Sciences of Physics, Mathematics, Biology, the Humanities and the emerging fields of Computer Science and Automata theory. The fascinating story of these meetings, the substantive discussions and the interpersonal issues that were played out, is told in the Publications of the Josiah Macy Foundation Symposia, in books by Heims (1991, Dupuy (1994), and in a section of Hayles (1999) "How we became posthuman". Transcripts of the proceedings were recently published by Pias (2003).

I consider these meetings the birthplace of the "Digital Brain" and origin of the associated influential notions of computation, representation, information, and the single neuron record. Their entrenchment in the Neuroscience discourse was initially quite explicit and in adherence to their original meaning. However, as time passed, their function in Neuroscience became less literal, but as metaphors more insidious. While the judicious use of metaphors can assist at times with intuitively illuminating a target domain (Arbib, 1972), they tend to carry with them the style of reasoning of the source domain which may be (and often is) quite inappropriate for the target; thus entailing the risk of tacitly contaminating the target with erroneous styles of reasoning.

The thrust of the following arguments is critical with a view of exposing conceptual flaws and neurobiologically unwarranted application of terms and ideas beyond their original sense, mostly from the physical sciences. This must not detract from the fact that the experimental work per se to which the investigators applied these notions, is frequently of the highest caliber, and deserving of admiration for ingenuity of experimental design and execution. My quarrel is with a single-minded interpretation of results and the discourse in which they are framed. For the investigators' jargon in 'the life of the laboratory', this is probably of lesser consequence, except for the price of excluding alternative theoretical frames which would have directed the research into different, possibly more revealing, channels.

It may be helpful to illustrate this point with a case study of multiple conceptual interpretations of the same kind of experimental data (Mountcastle and Werner, 1964; Werner and Mountcastle, 1963) recorded action potentials in single afferent nerve fibers and sensory neurons in response to precisely measured mechanical stimulation of peripheral receptors. In the first place, we determined the correlation function between stimulus strength and neural activity, the latter as frequency in bins of various duration. We then subjected the same data to analysis in three different conceptual frameworks: as a psychophysical (Weber) function, as a statistical decision process on distribution functions of static responses, and as information transmission from the stimulus to the neural response domain. What I am advocating with this example is the possibility, under suitable conditions, to interpret one and the same set of experimental observations in multiple theoretical frames of reference, each reflecting a different theory of nervous system function (Werner, 1985). This is an example of interdependence of theory and neural data (Hardcastle and Stewart, 2003).

The real problem arises with the study of Cognition and Consciousness when the Neuroscience reports are taken at their face value outside the laboratory of origin: it is at this point that unjustified and misleading generalizations can attribute meaning to the reports beyond