David W. Salt

A Truly Hard Problem

A Review of 'Neural Correlates of Consciousness' edited by Thomas Metzinger[1]

> Consciousness studies, like most areas of study, is replete with cottage industries and fads — *Valerie Gray Hardcastle* (p. 259).

This collection of papers presents a very wide range of findings both empirical and conceptual, as its sub-title suggests. It is undoubtedly an essential textbook for anyone interested in the study of consciousness whatever their background. Some of the papers in this collection are destined to become classic foundational studies in the emerging study of the neural correlates of consciousness. Others are less likely to stand the test of time but in the state of our current theories still make outstanding contributions to our understanding and interpretation of NCC research.

However, there are also inevitable problems with this collection which must be addressed in order to write an honest review. The first problem is that to fully understand every contribution to the volume would require a breadth of expertise that few of us possess. The empirical findings of neuroscientists stand alongside works of pure philosophy whilst agency and social cognition are also given coverage. Many of the contributions recognize the inherent problems, for instance that some working neuroscientists have little time for philosophizing preferring to get on with the empirical work. It is impossible to come to such a collection without bringing some personal baggage which will influence our viewpoints and great care needs to be exercised in criticising work not in our own primary subject areas.

A secondary, though related, problem is that sometimes you can't see the emperor's new clothes and sometimes you'd swear the authors have never heard of Ockham's razor. This is not meant to be a cheap jibe but a recognition of the difficulties inherent in assembling such a collection and the inevitable pitfalls of

Correspondence: David W. Salt, School of Mathematics & Computing, University of Derby, Kedleston Road, Derby DE22 1GB

[1] **Thomas Metzinger (ed)**, *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, Cambridge, MA and London: MIT Press, 350 pp., $50.00, £34.95, ISBN 0 262 13370 9 (hbk).

being brave enough to express an opinion at this stage of the game. Where you think the weaknesses lie in some of the reasoning in these articles will, I suspect, also be strongly influenced by your own background.

The book is organized into five sections; the first considers 'Foundational issues and conceptual problems'; sections two three and four consider 'Candidates for the neural correlates of consciousness' under groupings of representational dynamics, vision and 'Consciousness, anesthesia and the NMDA receptor complex'; the final section is entitled 'Toward the neural correlates of selfhood, agency and social cognition'.

It is easiest to start with part 3 of the book (Candidates for NCC: Vision) and work outwards, for the simple reason that the work presented in the chapters in this section describe the most obviously empirical work reported in the book. For instance Schmidt ('Visual perception without awareness: Priming responses by color') describes the experimental results which show some evidence for colour processing without awareness and discusses the results in a noncontroversial way. Kahuna ('Face representation without conscious processing') also follows the route of presenting detailed experimental work which could be easily reproduced and ends with a general discussion which interprets the results without wandering into the realms of speculation. Dominic ffytche ('Imaging conscious vision') ventures into conceptual areas with a modular model of consciousness but still concludes modestly: 'Visual consciousness correlates with some neural activities within areas specialized for the attribute perceived.'

Section 4 ('Candidates for NCC: Consciousness, anesthesia and the NMDA receptor complex') is similarly empirical in its approach and is centred on the first paper in the section by Flohr ('NMDA receptor-mediated computational processes and phenomenal consciousness'). Arguing that Higher Order Representations require integrated experience of objects and ourselves in the world he describes his hypothesis succinctly 'that the NMDA synapse implements the binding mechanism that the brain uses to produce these large-scale representations to which these higher-order processes belong' (p. 253). He demonstrates the role of the NMDA receptor complex by reference to the observed psychoactive effects of anaesthetic drugs.

The results reported by Flohr are potentially very exciting but, as Valerie Hardcastle points out in one of the most trenchant papers in the book ('How to understand the N in NCC'), showing that something is a candidate for NCC is a far cry from establishing its primacy. Although it could be the NMDA receptors, it could equally be the cell assembly itself (her own viewpoint), or the reticular formation, or internal quantum effects, and so on. (Interestingly only Hardcastle and Crick & Koch make reference to Penrose's theories in this volume.)

Working back to section 2 of the book ('Candidates for the NCC: Representational dynamics') we begin to run into more controversial areas. Crick & Koch posit an 'unconscious homunculus' who receives data through the senses and carries out actions. It is only a representation of the unconscious homunculus' activities that emerges into the conscious mind. They go on to narrow down the sense of self to 'a narrow band of neuronal activity between inner and outer

worlds' whilst admitting that how this produces the subjective world is 'still a complete mystery'.

Damasio ('A neurobiology for consciousness') takes a different tack in proposing a duality of core and extended consciousness resulting in a number of selves (core, autobiographical and protoself). Singer ('Phenomenal awareness from a neurobiological perspective') suggests that mental phenomena emerge from meta-representations. The implementation of these meta-representations are 'implemented by the dynamic association of distributed neurons into functionally coherent assemblies' (p. 134). Parallel and recursive signalling giving rise to 'reentry' in the brain are discussed by Edelman and Tonini and have some similarities with Singer's model in that they require recursive processing to achieve emergence of consciousness.

Section 1 considers the foundational and conceptual problems of the field. Revonsuo gives a good account of the metaphorical approaches to understanding consciousness, a rare approach when so many scientists seem unable to distinguish between metaphor and reality these days. However I was not ultimately convinced by his plan for a research programme for NCCs. Whilst I may have misunderstood his intentions, his paper read as though he wanted to define the programme to discourage dissent from the assumptions that he proposed. A case of wanting to have his cake and eat it too.

Although it may be unfair to the other authors in this collection, it seems to me that David Chalmers' paper was the centerpiece both of this section and of the book. With great clarity he picks his way through the philosophical issues to be addressed by the very basic question of 'What is a neural correlate of consciousness?' The very fact that answering this question is so difficult shows the care with which the subject should be approached. Whilst it is necessary to read his piece in its entirety to appreciate the conclusions I nevertheless think that it is useful to report his two overall definitions :

- An NCC is a minimal neural system N such that there is a mapping from states of N to states of consciousness, where a given state of N is sufficient, under conditions C, for the corresponding state of consciousness.
- An NCC (for content) is a minimal neural representation system N such that representations of a content in N is sufficient, under conditions C, for representation of that content in consciousness (p. 31).

The beauty of Chalmers' piece, though, is the precise and logical way that he takes the reader though the options and the problems involved in arriving at these two definitions. Even after arriving at these definitions he has to admit 'Of course the precise nature of condition C is still debatable' (p. 31). Nevertheless Chalmers has once more cut through a confusing subject area and highlighted what can be precisely articulated about the subject at this point in time and what remains to be fully worked out.

I have one criticism of Chalmers' piece and that is that he makes frequent reference to the systematicity of brain processes but not directly to systemic properties. I believe that there is still a gap to be filled here and that many of the

empirical papers in this collection point to a need for systemic analyses that are different in kind to systematic analyses. This is not a trivial point because producing a step by step description (systematicity) may give an entirely different account of events from a systemic interpretation of the behaviour under scrutiny. The difficulty is compounded by the fact that systemic interpretations may lean too heavily on vague notions of emergence which may be why so many authors in this volume avoided more holistic approaches.

I have left the final section on the NCs of selfhood, Agency and Social Cognition until last and I will confess that I found this section least satisfying, though I must attribute this largely to my own background and knowledge. Metzinger gives a good account of his representationalist theory of subjective experience, Proust compares two current approaches to schizophrenia from a philosophical perspective and Gallese explores links between the world of subjective experiences and social cognition. All are clearly expressed and well constructed but I personally found them hard to correlate with the proceeding four sections. I think that part of the problem is that the book lacks an afterword or summation which could have been used to contextualise each section in relation to each other and to fill in some gaps for readers with differing backgrounds. Whilst Metzinger addresses these concerns partially in his introduction I still felt the need for the strands to be drawn together at the end of the volume with some pointers towards future directions.

As I indicated at the outset it is almost impossible to do full justice to this collection without an exceedingly broad and deep understanding of several interrelated disciplines. Indeed it is because of this huge range of subjects that this book is so welcome. Researchers in many different fields have become aware of the central importance of NCCs and are hungry for an exchange of ideas that will inspire them to continue their work and to forge an inter-disciplinary approach.

Inevitably I will have misunderstood or misinterpreted some of the work I have reviewed here but I welcome each and every one of the articles in this book and warmly recommend it to all readers of this journal. Anyone who sticks their head above the parapet at this point deserves to be given a fair hearing and I hope this principle can be extended to the reviewer. I'd like to give the last, as well as the first, word to Valerie Hardcastle. Whilst this book is an excellent collection we still have a long way to go.

> For now, we have only educated guesses, personal declarations of faith, and a plethora of individual research programs. But much basic research remains to be done and, more important for our concerns, our fundamental theoretical scaffolding remains to be constructed. For now, the NCC remains a truly Hard Problem with no solution in sight (p. 264).