Cognitive Models Go Paddling in the Waters of Consciousness:

Review of *Scientific Approaches To Consciousness*, Jonathan Cohen & Jonathan Schooler (Eds.)

Mark C. Price
Cognitive Section
Institutt for Samfunnspsykologi
University of Bergen, Bergen
NORWAY

mark.price@psych.uib.no

PSYCHE, 5(15), July, 1999
http://psyche.cs.monash.edu.au/v5/psyche-5-15-price.html

KEYWORDS: attention, cognition, connectionism, first person data, functionalism, hard problem, implicit learning, implicit memory, metacognition, neural activation, subliminal perception.

REVIEW OF: J.D. Cohen & J.W. Schooler (Eds.), *Scientific Approaches to Consciousness.* Mahwah, NJ: Lawrence Erlbaum Associates. xii +538pp. ISBN: 0-8058-1472-8. Price $45.

1. Overview

Scientific Approaches to Consciousness is a collection of 28 chapters based on the 25th Carnegie Mellon Symposium on Cognition. The chapters, which vary considerably in quality and depth, provide a cross section of empirical work and theoretical positions from many well known researchers within the cognitivist/functionalist tradition. Readers will be disappointed if they expect a broad interdisciplinary approach to consciousness, or a blinding insight into the "Hard Problem" of wet consciousness. And many chapters will prove hard going for the non-specialist in the cognitive brain sciences. Little of the content covers new work, but for advanced students, researchers or teachers, the book can provide a useful overview of selected research areas outside their immediate area of specialisation. It also provides a sample of what some of the big names in the cognitive brain sciences have to say about consciousness.

The editors set the cognitivist/functionalist tone for the book in their introduction. As a criterion for consciousness they propose a "Cartesian Turing Test" - if the behaviour of a system is indistinguishable from that of a human being, the system is conscious. They state that the aim of the book is to ask whether "Is it now possible to study consciousness directly in a scientific way?" (p. 5). The reader might therefore be forgiven for expecting some discussion of what counts as "science". However such discussion is avoided by most contributors. Even in the introduction we are left only with the brief statement that science is limited to the domain of the physical and observable, and does not encompass "philosophical" issues such as the dualism debate. This conception of science will seem overly narrow to those who consider that our notion of the physical is constantly evolving, or that constructs can have explanatory value in science even if they are not direct observables, or that the issue of whether dualism is needed to explain consciousness is an integral part of the scientific enterprise.

The authors were apparently asked to "let their hair down" but discussions of consciousness in this book contain nothing of the radical speculative reappraisals of the scientific world view that appear in recent edited volumes such as Jonathan Shear's *Explaining Consciousness: The Hard Problem* (Shear, 1997). Chalmer's "Hard Problem" of how brains could conceivably give rise to the phenomenon of subjective first person consciousness (Chalmers, 1995) is largely ignored, although a few authors, such as neuropsychologist Martha Farah are at least honest enough to explicitly admit this. Philosopher David Galin remarks in his short chapter that in the symposium on which the book is based, "there was enormous resistance to talking about subjective experience" (p. 447). This is reflected in the overall content of the book, although to be fair a few of the chapters do address subjective experience directly.

The chapters are organised into 6 main sections: Attention and Automaticity, Subliminal Perception, Implicit Learning and Memory, Metacognitive Processes, Neuropsychological and Neurobiological Approaches, and Theoretical Issues and Approaches. The way some chapters are fitted into this organisation sometimes seems rather arbitrary, although this is an understandable difficulty with themes that can be multiply interlinked. However the structure of the book could have been much more integrated if there had been a greater attempt at editorial synthesis. The few short linkage-chapters, presumably included with this aim, are for the most part disappointing and provide little more synthesis than a few extra comments in the introduction would have done. In order to better tease out how the various chapters contribute to an understanding of consciousness, the following discussion of individual chapters therefore departs from the book groupings listed above.

2. Implicit Cognition

Central to many contributions in the book is the idea that one way to progress in studying consciousness is to study nonconscious aspects of cognition, and examine how these contrast with more conscious aspects.

Spread among the first 3 sections (Attention and Automaticity, Subliminal Perception, and Implicit Learning) are contributions from researchers who have been providing some of the most significant methodological innovations in the study of nonconscious processing, including the areas of nonconscious, or "implicit" perception, learning and memory. These processes are referred to as "implicit" since they are implicit in behaviour, rather than explicit in consciousness.

Separate chapters by cognitive experimental psychologists Larry Jacoby, Philip Merikle and Anthony Greenwald, together with their co-workers, all begin by outlining the methodological problems that have plagued attempts to show measurable nonconscious processing of information that is not consciously represented. The problems centre on the difficulties of (1) finding experimental measures that exclusively or exhaustively reflect either just nonconscious processing or just the contents of conscious representations, and (2) establishing with any degree of statistical reliability that task-relevant conscious representations are completely absent. Since solutions to these problems must address the question of how we can operationally define and measure conscious awareness, they are widely relevant to the whole issue of mapping the relationship between consciousness and cognition.

Jacoby et al. provide a useful summary of one solution to the above problems: the "process-dissociation procedure", which claims to quantitatively estimate the additive but separate contributions that conscious and nonconscious processes make to performance on a single task. This is achieved by exploiting the qualitative differences between consciously controlled processes, defined as "those producing a difference in performance that reflects current intentions.", and unconscious processes, defined as "producing the same effect regardless of whether the effect is in concert with or in opposition to current intentions." (p. 21). Pairs of experiments are then run in which conscious and nonconscious processes are pitted either against each other, or alongside each other, so that the contribution of each type of process can be calculated like variables in a pair of simultaneous equations. A good overview is provided of the wide applicability of the method, including in clinical work. Some converging evidence is given for the important assumption that conscious and nonconscious processes are (at least under certain conditions) truly independent, although there is no extended discussion of the arguments against this controversial claim.

Greenwald and Draine's chapter is more of an experimental paper than a research overview. Detailed description is given of 3 experiments (one previously published; Greenwald, Klinger, & Schuh, 1995) which claim to show that the meaning of words on a computer display, which are masked so they cannot be consciously detected or located, can nevertheless bias "guessed" judgements about the words' locations. The paradigm, and the statistical method used to avoid the problems of establishing null sensitivity on a measure of conscious awareness, are both ingenious (though not without their critics). However a meta-analysis over several hundred subjects is needed to support the existence of a very small nonconscious bias, so the authors' conclusion that subliminal techniques could be used in the mass media to "produce significant influences on behavior" (p. 106) seems premature to say the least.

Like Jacoby et al., Merikle and Joordens redirect the debate away from the issue of whether stimuli have tiny measurable effects when conscious representations are absent, and concentrate on the more interesting question of whether consciously and nonconsciously perceived stimuli have qualitatively different effects on behaviour. Merikle and Joordens take up the Jacoby et al. idea that nonconsciously perceived stimuli can affect behaviour in an automatic way which contradicts conscious task instructions; e.g. nonconsciously presented words (TABLE) are likely to be used to complete subsequent word stems (TAB_ _), despite the instruction NOT to complete the stem with the preceding word. Further evidence that performance on this kind of "exclusion task" is a useful operational measure of conscious vs. nonconscious processing is then provided by experiments which show it converges with other qualitative differences in behaviour; e.g. with the ability to use the predictive contingencies between stimuli to influence the "Stroop" priming effect that colour-words have on colour-patch naming.

One important area of implicit cognition centres on the study of implicit learning, defined in Arthur Reber's chapter as the process "whereby reasonably complex knowledge about structured stimulus displays is acquired largely independently of awareness of both the process and the products of acquisition." (p. 137).

Pawel Lewicki and co-workers provide a nice review of some of the experimental paradigms used in this research. Their design is often expanded on in footnotes which will help the non-specialist reader to follow the main text. This is an editing strategy which could have been usefully employed in many of the other chapters. Lewicki et al.'s chapter also stresses the role of implicit learning mechanisms in the wider context of everyday learning (including acquisition of expert knowledge, and the development of self-perpetuating biases).

Reber's chapter takes a certain familiarity with the basic paradigms for granted, and concentrates on arguing against criticisms that subjects in so-called implicit learning experiments may after-all be conscious of what they are learning. Some of his arguments centre on distinguishing between different shades of consciousness. Subjects may have feelings of correctness or confidence in their responses, but although these may be conscious, they do not constitute consciousness of the rules that are being learned. These arguments emphasise the importance of developing a more sophisticated analysis of the contents of consciousness, and thereby avoiding some of the polarised debates as to what counts as conscious or nonconscious processing. There is an interesting (but unmentioned) link here with work on what has been termed "fringe" consciousness by authors such as Mangan (1993). Reber's second main line of argument is to focus on the many qualitative differences between implicit and explicit processes; e.g. implicit processes are more resistant to injury, less sensitive to individual differences or affective state, and less sensitive to interference from concurrent tasks. Reber concludes that consciousness is related to the explicit top-down modulatory aspect of cognition, which interacts synergistically with aspects of the functioning of an evolutionarily more "primitive" and more parallel implicit system. This implicit system, which in humans can pick up remote, complex and multiconditional patterns, is an extension the mechanisms which evolved to detect simple environmental contingencies in lower animals.

Donelson Dulany's contribution to the implicit learning section is stylistically dense and very heavy going. He presents a provoking attack on a simple equation of implicit learning with nonconscious processes and explicit learning with conscious processes. One strand to this attack is a summary of the many methodological criticisms that have been levelled against the kinds of implicit learning paradigms described by Reber and Lewicki et al. A claimed catalogue of artefacts, non-replication and poor assessment of conscious awareness lead him to conclude no learning is found outside of consciousness. For the reader, comparison with the preceding chapters is like listening to politicians discussing economic statistics; without a sabbatical to dig through the empirical literature, it is difficult to tell who is correct. The second strand of Dulany's attack is that implicit and explicit learning do not differ with respect to being conscious or unconscious, but with respect to the types of conscious content they are associated with (e.g. non-propositional versus propositional). Here of course there is much convergence with Reber's chapter. But whereas Reber is a fan of the cognitive unconscious, and stresses the extent to which our behaviour is guided by aspects of knowledge that are nonconsciously represented, Dulany has a different philosophical axe to grind. He emphasises the conscious aspects of all learning situations, mounts a disparaging tirade against the whole notion of a cognitive unconscious which merely attributes "the poorly understood to a residual humanoid that secretly does what persons do", and concludes "there's nobody home but us." (p. 206). As Reber points out, Dulany's position therefore exemplifies the influence of Cartesian and Lockean philosophical traditions for whom the mind is a conscious mind, transparent unto itself.

3. What Information-Processing Conditions are Necessary for Consciousness?

While providing a useful overview of methodological advances in disentangling conscious from nonconscious processing, the preceding chapters do not attempt to discuss the theoretical implications of their empirical work for understanding WHY certain operationally defined processing styles happen to be accompanied by consciousness. This task is left to other chapters, which are again rather dispersed between the different sections of the book.

Lynne Reder and Jolene Scully Gordon's chapter is one of many that propose consciousness is related to high levels of neural activation. They claim that the qualitatively different effects of conscious and nonconscious stimuli can easily be accounted for by current cognitive models. One such model (complete with undefined acronym) merely proposes that low activation levels can allow effects such as automatic priming to occur, but fail to cross the "threshold" into consciousness or engage the production rules that allow information to be used in accordance with task instructions.

For Marcel Kinsbourne, the importance of activation is related to the idea that the contents of a representation become conscious if they are "integrated into the *dominant

focus* of neural activity - that is, the predominant pattern of neuronal firing constituting the brain state (network attractor state), which is in current control of response processes, including verbal and imaged thought and expression." (p. 343). Eligibility for such integration is partly determined by the duration and level of activation of a representation. However implicit processing may again still occur if activation falls short of the requirements for consciousness.

As evidence for the importance of activation, Kinsbourne notes that the severity of neuropsychological deficits that involve impairments of conscious awareness is not always absolute, but can be dependent on temporarily induced states. For example, deficits in attending one side of space (unilateral attentional neglect) that are caused by damage to one cortical hemisphere can be temporarily remediated by hemispheric activation such as pouring water into a patient's ear ("caloric stimulation"). Kinsbourne argues that if a state change can have such an effect, the neural substrates of the "impaired" representations cannot have been completely destroyed, but rather must merely have been underactivated. Unfortunately there is no dialogue here with Rumelhart's later chapter which outlines a connectionist account of how alternative state-specific versions of a stored representation can co-exist and hence presumably suffer selective damage. It is also a shame that fascinating impairments of consciousness like attentional neglect, which deserve chapters to themselves, are not discussed in more detail in this book.

Other interesting speculations are introduced but unfortunately hurried over by Kinsbourne, such as the idea that certain neuropsychological deficits result in consciousness of stages of processing that are normally conscious for too short a period to be remembered and integrated into our main stream of conscious. Or the idea that parallel consciousnesses may reside within non-damaged brains, so that supposedly "nonconsciously" mediated processing may be nonconscious only from the point of view of the verbal report system.

Taking a slightly different tack, Martha Farah et al. provide a clear and easily digested summary of their idea that consciousness is impaired when representations are degraded, rather than when they are disconnected from a dedicated "consciousness subsystem" or, as suggested by the preceding authors, are merely underactivated. This chapter manages to be concise without being hurried, and has the additional merit of being grounded in working computational models. Farah et al. focus on the neuropsychological deficit of prosopagnosia, in which brain damaged patients are severely impaired in explicitly/consciously recognising faces, but can nevertheless show behavioural evidence of a residual implicit/nonconscious ability to recognise faces. Farah et al. then describe how this dissociation between impaired explicit recognition and preserved indirect effects of implicit recognition can be simulated by "lesioning"a connectionist network model of face recognition. The authors are refreshingly honest enough to admit that although they are able to throw light on why explicit and implicit measures dissociate, none of this explains why consciousness should be associated with explicit measures in the first place.

Another attempt to argue that conscious representations are more than just a highly activated version of unconscious representations is provided by George Mandler's chapter. Mandler's analysis procedes at a higher level of description than that of Farah et al., arguing that conscious representations are *constructed* from activated unconscious schemata. This analysis of the qualitative differences between conscious and nonconscious representations draws heavily and openly on the ideas of Marcel (1983), but does not go into enough detail to do these ideas justice or make them clear to the reader. According to Mandler, activation levels and consciousness are still linked in the different sense that the nature of the conscious representations gives rise to extra activation which enhances their cognitive accessibility and thereby plays a crucial role in executive control and selection processes. Here there is an obvious, but undiscussed, link with the chapters on consciousness and attention, and with Baar's Global Workspace theory of consciousness (see below).

John Kihlstrom leaves activation levels out of the picture altogether, and instead concentrates on his proposal that representations enter consciousness by virtue of being linked, within Working Memory (non-psychologists can take the latter concept as roughly equivalent to a modern reformulation of the idea of active short-term memory which incorporates a cluster of creative "information-manipulating" processes), to some kind of currently active representation of "self". It is argued that general cognitive models such as ACT are able to accommodate such an idea. The distinction between explicit and implicit cognition is interpreted with respect to this linkage with "self" representations. For example, if the link is not made at all, implicit/nonconscious perception may still occur. If the link is made, but not stored in long-term memory, implicit memory may still occur. This chapter is not a bad summary of Kihlstrom's position, is quite readable, and extends the discussion to interesting areas such as multiple personality disorder which is attributed to the occurrence of parallel concurrent self representations. However the crucial question of the exact nature of the proposed "link" between a self representation and other representations is left unexplained, as is the reason why this link should give rise to phenomenal experience.

Kihlstrom's hypothesis does not necessarily contradict the idea that other factors such as activation levels are also important for consciousness. So much is hinted at in J. Allen Hobson's chapter (of which more later) where studies of sleep vs. waking consciousness are taken to support the importance of activation level for consciousness, but where it is also argued that "Consciousness is a graded, global integration of multiple cognitive functions yielding a unified representation of the of the world, our bodies, and our selves." (p. 395).

A greater attempt by the editors to tie together the themes introduced in these various chapters would have been very useful. However the chapters are geographically dispersed around the book, and the reader is given little help in resolving where the themes are in contradiction, where they are complementary to each other, or where they represent analyses at different levels of description. An exception is Jonathan Cohen's short chapter which aims to resolve a claimed "apparent conflict" between Kinsbourne's and Hobson's chapters that was in fact not particularly apparent to this reader. Much of Cohen's chapter

is unhelpfully spent merely summarising Kinsbourne and Hobson, and is an example of a commentary chapter that adds little of substance.

4. The Relation of Attention and Executive Processes to Consciousness

At a folk psychological level, there is great overlap between the concept of consciousness and that of attention. However a simple mapping between (a) the conscious-nonconscious dimension and (b) the common conceptual distinction between "attentional" and "automatic" processing, is challenged in a chapter by Richard Shiffrin. He argues that there is dissociation not only among the various operational criteria traditionally used to distinguish "automatic" from "attentional" processing (e.g. control, speed, unlearning rate, serial vs. parallel processing etc.), but that these two categories do not always map respectively against nonconscious and conscious processing. While it is good to have one of the godfathers of the automatic/attentional distinction pulling apart this text book mapping (which is explicitly adopted in the chapter of Jacoby et al.), Shiffrin's chapter is very dense and will prove difficult for the non-specialist reader without a working knowledge of the many experimental paradigms that are introduced by name but not described. The conclusion of the chapter also rather leaves the reader hanging in mid-air: having argued that a clear relation between automatic/attentional and nonconscious/conscious is unsupported, he maintains that there remains a strong positive relation between consciousness and attention which deserves further exploration. However no attempt is made to resolve this interesting and crucial apparent contradiction.

An additional source of frustration with Shiffrin's chapter is an often insufficient regard for phenomenology. For example he claims that the automisation of cognitive processes is not always reflected in changes in the contents of consciousness. However automisation is known to produce some very salient differences in phenomenology, such as the difference between serial visual search and the "pop-out" that accompanies parallel search, and one suspects that Shiffrin's counter-examples may stem at least partly from our currently impoverished methods of measuring and classifying subjective experience. But to his credit Shiffrin is prepared to challenge at least one commonly-held phenomenological assumption which is explicitly stated by Herbert Simon in a later chapter, that the seriality of higher cognition, as reflected in consciousness, is unquestionable. Shiffrin makes the interesting statement that although consciousness sometimes appears very serial (as with linguistically mediated thoughts), it is not evident how one establishes this. Again the obvious unstated demand is for the evolution of a better methodology for charting the contents of consciousness. Only a few of the contributors to this volume are concerned with taking up this challenge.

The accompanying chapter by Walter Schneider and Mark Pimm-Smith typifies this lack of concern with phenomenology. A stereotypically functionalist model of consciousness is described, with consciousness mapped to a subset of high level executive control

processes referred to as the "message aware control system". In the course of the chapter, the authors explicitly propose that subjects taking part in a varied-mapping visual search task are unaware of the way their attention serially switches between the items in the display and the target items they are holding in their memory. But again, how adequately has the phenomenology that accompanies such experimental tasks really been tested? In similar vein, Schneider and Pimm-Smith's general mapping of consciousness to a particular aspect of attentional control does not seem to do justice to the phenomenological variety and richness of consciousness. Like Shiffrin's chapter, this chapter is too dense to be useful to the non-specialist. However it is unlikely to provide much of substance to the specialist reader either, and there are several places where the text is so badly written as to be almost impenetrable. In addition, the rather homuncular construct onto which consciousness is grafted is given no empirical justification and seems derived from the kind of unitary "central controller" conception of attention which some will find a little outdated.

Mandler's chapter, referred to already above, maps consciousness to an attentional bottleneck in an information processing system, and therefore gives it a similar role to the one proposed by Schneider and Pimm-Smith. But at least Mandler gives the reader a good outline of why the architecture of complex system like the human brain requires a serial limited-capacity selective bottleneck in order to successfully map multiple inputs and vast numbers of parallel unconscious processes to a coherent behavioural output. Needless to say, the explanatory role of consciousness in this bottleneck is left ambiguous.

A chapter by Bernard Baars and colleagues offers much the best attempt in the book to map consciousness to aspects of executive control. Baars' "contrastive analysis" approach is outlined, in which meta-analysis of the experimental literature is used to draw up a list of the differences between conscious and unconscious processes that must constrain any theory of consciousness.

Here there is an obvious thematic link with the implicit cognition chapters. There is also an interesting contradiction with Shiffrin's chapter, since Baars' list closely resembles the list of variables traditionally used to define and distinguish automatic from attentional processes. In contrast with Shiffrin, Baars is therefore attaching more importance to the general pattern of qualitative differences between conscious and nonconscious processing, than to the exceptions. Some editorial discussion would have been useful here.

The data of "contrastive analysis" leads Baars et al. to Baars' now well-known Global Workspace theory, in which consciousness is seen as the spotlight of attention in Working Memory. This spotlight plays a mediating role in directing external (perceptual) or internal input to nonconscious processing subsystems. Consciousness is thus linked to the "global access" and dissemination and co-ordination of information. The focus of the chapter is not however Global Workspace theory itself, but the way in which the theory can be used to explain "absent minded" action slips in terms of a momentary loss of control by the limited-capacity spotlight of attention. A good summary is provided of

some very neat empirical lab work on induced action slips. Although Baars has often argued that he is not proposing a straight equation of consciousness with Working Memory or with attention or with executive systems, the inescapable drift of this chapter is indeed that consciousness is being mapped to a function that corresponds to a subset of attention and executive control.

5. First Person Data

The phenomenological limitations of some of the attention chapters, and the debates raised in the implicit learning chapters, both point to the importance of more sophisticated analyses of the contents of consciousness. Two chapters provide a therefore welcome attempt to take first-person data seriously, and demonstrate that collecting such data can be very much within the fold of the scientific method.

Suparna Rajaram and Henry Roediger provide a very readable introduction to research on one type of meta-memory judgement: the subjective distinction between episodically "remembering" a test item and simply "knowing" you have seen it before. They describe how these different subjective states are differentially affected by variables which overlap those known to differentially affect implicit versus explicit memory measures. This chapter is honest about many questions that have yet to be adequately addressed (e.g. the need for more precise definition of what exactly it means to say a subject "knows" without "remembering"), but is brave enough to begin the journey. Another chapter, by Jonathan Schooler and Stephen Fiore, concentrates on the extent to which the contents of consciousness are not always easy to verbalise. While this is an important point, the empirical examples used to illustrate non-verbal consciousness are not always the best. Moreover, although the authors argue that self-report measures are needed to reveal non-verbal consciousness, they miss the chance to discuss and resolve the apparent contradiction between using (verbal) self-report to indicate non-verbal consciousness. However they do provide some interesting examples of how verbalisation can actually impair performance in a number of experimental paradigms.

These two chapters stand in marked contrast to misgivings, raised by some of the other authors, about the scientific value of experiential reports. For example, Clark Gylmour brings up the weary jaded argument about the essential privacy of consciousness. Kinsbourne brings up the more useful argument that past subjective states can never be verified because recollecting them is always tantamount to the recreation of new subjective states. Hence the relation between the physical and the subjective can never be rendered transparent. Unfortunately Kinsbourne then contradicts himself by proposing that the task of neuropsychologists is to determine the nature of the neural states that accompany subjective states; surely this task is rendered impossible if subjective states cannot ever be reliably measured!

6. Accounting for the Variety of Contents of Consciousness

Continuing with the theme that we must play close attention to first- person data, three chapters are concerned with the task of relating the wide variety in the content and quality of our conscious representations to underlying brain events.

Daniel Wegner's particularly well written and interesting chapter asks why our conscious thoughts are constantly wandering. His proposed explanation rests on a theory of mental control, in which an operating process and a test/monitoring process act against each other in a tug-of-war. The wandering of the mind towards distracters is seen as a normal consequence of the control system, and the theory claims to account for interesting experimental data on how, under high load conditions, trying to suppress information can actually bring it to mind (e.g., as in Freudian slips).

Marcia Johnson and John Reeder try to account for a much broader range of conscious contents in terms of one attempt to model global cognitive architecture at an "intermediate" level of description. Their particular point is that the wide variety of possible contents is best explained as a consequence of which subsystems are, at any given moment, monitoring and controlling which other subsystems. The authors contrast this information-processing view of what determines whether particular representations make it into consciousness, with some of the others summarised in section 3 above. Their approach can also be linked to the chapters on attention and executive control.

Hobson's chapter tackles the state dependency of our contents of consciousness, with particular respect to the differences in these contents during waking states, REM sleep and NREM sleep. Hobson gives a useful summary of experimental evidence that the phenomenological differences between these states can be accounted for in terms of three variables: neural activation level, neural input-output gating, and chemical modulation of neural transmission. A great merit of this work is that it relies on, and brings together, data from both experiential and neural levels of description. On the other hand it is a slight shame that the book contains no reports of work on lucid dreaming to counter Hobson's assertion that voluntary control and self-reflexive awareness are always absent during REM sleep.

7. Consciousness and Connectionism

Connectionist, neural-network models of brain functioning continue to play an increasing role in our understanding of the mind. It is therefore appropriate that two of the big names in connectionist modelling should have contributed to this book. However the contributions are disappointing, and less illuminating than the practical applications of such models found in chapters such as Farah's.

A chapter by David Rumelhart begins with a confused discussion of the mind-body problem, which is confounded with the relationship between system hardware and software. The chapter then slips into a brief outline of how connectionist networks can model certain very broad aspects of state dependent learning and mood congruency. There is no attempt to tie this in with consciousness, no conclusions are provided, and the account includes completely opaque one-liners about how networks settle into "good" interpretations of input data (with equally opaque equations thrown in to show this is real science). The chapter is a classic straw man for those who argue that scientists claiming to write on consciousness often end up writing about something else.

James McClelland then paints an overview of the mind as a connectionist system in which implicit memories result from altered connection strengths within a network, whereas explicit memories occur when original cortical activation patterns are reconstructed via interconnections with stored activation patterns in the hippocampus. The relationships between this view and those expressed by some of the other authors are discussed. For example, Kihlstrom's view that consciousness is tied to representations of self is given a rough ride. McClelland also manages to extrapolate unconvincingly from Rumelhart's account of state dependency to the idea that connectionist models help us to explain emotions.

8. Philosophical and Metaphysical Issues

Some of the broader conceptual issues concerning the framework for consciousness research are explicitly addressed in a couple of excellent chapters by philosophers.

Owen Flanagan argues that a unified theory of consciousness may be possible even if consciousness has heterogenous contents or origins; the idea that homogeneity is NOT a requirement for theoretical unity is illustrated by analogy with other areas of science. Flanagan also argues for his "natural method", in which phenomenological, cognitive and neurophysiological approaches are all respected. As an example, he addresses the argument over whether dreams are conscious experiences that actually occur during sleep, or (a la Dennett) are merely the experience of having had an experience. In this case it turns out that neurophysiological evidence, such as the occurrence of 40Hz oscillatory activity during REM sleep, supports the phenomenological impression that the consciousness really does occur during sleep.

David Galin's short but very entertaining chapter (entitled "What is the difference between a duck?") is another good reminder that philosophers still have a useful role to play in the consciousness debate. Galin encourages us to pause and ask why scientific attempts to explain consciousness are, for many people, always so unsatisfactory. He then poses the crucial question of what makes any explanation satisfying, and suggests that "An explanation is an account of the variance of the phenomenon of interest. It is satisfying if it is expressed in the terms of a frame of reference in which you are comfortable, and if it accounts for enough of the variance *for your purposes.*"(p. 449).

Galin then argues that most explanations of consciousness are unsatisfying precisely because they do not illuminate the issues that many people are interested in, such as the "wholeness" that is so characteristic of consciousness experience. Only when we start to ask about what we are most interested will our answers be "aligned with our deepest purposes" (p. 450).

The last chapter of the book also deals with more general issues, but is written by Herbert Simon of computational modelling fame, rather than by a philosopher. Simon strongly criticises the hand-waving generality of many of the cognitive models referred to in the book, and urges more detailed computer models whose predictions can be empirically tested. This is a complaint echoed by Gylmour who argues that it is unclear which kinds of data can distinguish between various cognitive models of consciousness. Simon's chapter also contains an interesting overview of serial versus parallel aspects of cognition, and concludes that research on consciousness will play a big role in furthering our understanding of this type of distinction. Quite how this will in turn increase our understanding of consciousness is not stated.

As is made clear in the book's introduction, metaphysical issues about consciousness are not part of the book's main remit. Although most authors are therefore obediently silent on such matters, a few are prepared to express some kind of view. For the most part these excursions are brief, unenlightening, and even downright confusing.

Regarding the ontological status of consciousness, McClelland proposes without much elaboration that consciousness is a "manifestation of certain properties of overall system function". A couple of authors are philosophically more specific. Kinsbourne seems to espouse a weak identity theory, in which consciousness IS neural activity: "...subjectivity is not a product of neuronal activity. It is the activity of the circuitry itself...it is what it is like for a neuron to be in a particular state...the human brain does not generate or produce consciousness...nothing additional is produced..." (p. 338). Something similar is proposed by Hobson who sees "the possibility of an ultimate identity between our conscious states as we perceive them and the unobservable states of our brain in which they arise" (p. 395).

Regarding the related issue of the causal status of consciousness, the closest we get to non-epiphenomenalism is in Johnson and Reeder's chapter. In a frustratingly short paragraph we are told that the function of consciousness with a big C may lie in its motivational value. However this potentially interesting theme is not fully explained or developed. Mandler is the only other author who seems prepared to talk as if consciousness "itself" was doing something. For example, discussing consciousness in the context of high-level selection, Mandler proposes: "A conscious state is constructed that reflects those aspects of the current search that do (partially and often inadequately) respond to the goal of the search. Consciousness at these points depicts waystations toward solutions and serves to restrict and focus subsequent pathways by selectively activating those currently within the conscious construction...The result is a directional flow of activation that would not have happened without the extra boost derived from the conscious state."(p. 488).

But does Mandler really mean to imply that consciousness has causal efficacy over and above the neural states from which it emerges? Does consciousness cause the extra activation, or does the activation cause the consciousness, or IS consciousness the extra activation? Of the other authors, at least McClelland interprets Mandler as describing a causal role for consciousness, and contrasts this to his own functionalist approach. But without further metaphysical precision, Mandler's way of speaking is ambiguous. The grass is always greener on the other side of the fence: When reading philosophers, one often wants to be rid of the waffle and get down to some real data; and when reading the scientists, one hankers for some philosophical rigour! The relative merits of dualism versus identity theory, or epiphenomenalism versus non-epiphenomenalism, may be beyond current empirical or conceptual resolution, but authors who do not put their (albeit agnostic) position plainly on the table leave themselves open to misinterpretation.

9. Conclusion

The book's introduction contains a call for consciousness to be addressed at different levels of analysis, and with different research approaches. However the range of the chapters does not meet this call. The volume is heavily biased towards work on implicit cognition, and towards attempts to relate consciousness to fairly global aspects of cognitive architecture. But even if the volume were re-titled "Consciousness in the cognitive sciences", thus permitting the exclusion of exotic cross-disciplinary contributions from areas such as quantum theory, many eligible topics would remain unrepresented. For example there are no chapters on the binding problem and neural oscillation, nor on the temporal microgenesis of consciousness, working memory, blindsight, attentional neglect, dorsal versus ventral stream processing, or altered states of consciousness to name but a few.

As a compilation, this volume can provide a useful taster of work in certain areas, but as a book it lacks discussion of the testability, mutual exclusivity or general distinction between the various models proposed. Opinion is bound to vary as to whether the contents dramatically further our understanding of consciousness. At the end of Flanagan's chapter the eternal million dollar questions are pushed aside and we are advised to "get on with the hard work of providing the right fine-grained analysis of conscious mental life and to see where it leads" (p. 421). For all the failings of the volume as an integrated work or a broad coverage, and for all the failings of individual chapters, many good examples are provided of the hard work that Flanagan advocates. Some readers will be sympathetic to Galin's skepticism over whether such work brings us much closer to a satisfactory explanation of consciousness. Others will side with Flanagan, grit their teeth, and keep walking.

References

Chalmers, D.J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2(3),* 200-219.

Greenwald, A.G., Klinger, M.R., & Schuh, E.S. (1995). Activation by marginally perceptible ('subliminal') stimuli: Dissociation of unconscious from conscious cognition. *Journal of Experimental Psychology: General, 124 (1),* 22-42.

Mangan, B. (1993). Taking phenomenology seriously: The 'fringe' and it's implications for cognitive research. *Consciousness and Cognition, 2(2),* 89-108.

Marcel, A.J. (1983). Conscious and unconscious perception: an approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology, 15,* 238-300.

Shear, J. (ed.) (1997). *Explaining consciousness: The hard problem.* Cambridge, MA: MIT Press.