

Learning to Work Together:

Review of *Consciousness In Philosophy And Cognitive Neuroscience* Antti Revonsuo & Matti Kamppinen (Eds.)

W.F.G. Haselager
Theoretical Psychology
Vrije Universiteit
Amsterdam
The Netherlands

also at

Cognitive Science Department
Katholieke Universiteit Nijmegen
The Netherlands

wfg.haselager@psy.vu.nl

Copyright (c) W.F.G. Haselager 1998

PSYCHE, 4(9), June 1998
<http://psyche.cs.monash.edu.au/v4/psyche-4-09-haselager.html>

KEYWORDS: consciousness, philosophy, neuropsychology, neurobiology.

REVIEW OF: Antti Revonsuo & Matti Kamppinen (Eds.). (1994). *Consciousness in philosophy and cognitive neuroscience*. Hillsdale: Lawrence Erlbaum, xii + 298pp. ISBN: 0-8058-1509-0. Price: \$US59.95 hbk.

This book is based on papers presented at the International Symposium of Consciousness held at the University of Turku, Finland, in 1992. As the editors claim, the explicit aim of the conference was to relate philosophical views with empirical findings by encouraging the dialogue between philosophers and cognitive scientists. The editors and participants (Bernard Baars, Patricia Smith Churchland, Daniel Dennett, John Haugeland, Vilayanur Ramachandran, Antti Revonsuo, John Searle, and Andrew Young, among others) have done an admirable job in presenting more than a loose collection of essays. In contrast to many edited volumes based on conference proceedings, the authors actually do address positions defended by other contributors, thus contributing to the coherence of the book.

Books of this kind may not, in general, be read first page to last, but this one contains valuable material to choose from.

The book is composed of four thematic parts, dealing with the ontological status of consciousness, its problematic aspects, its models, and its future. Each part is preceded by a short introduction by Revonsuo, outlining how the papers relate to the specific theme. The general introduction by Revonsuo and Sajama provides a brief review of the history of the problem of consciousness in philosophy and they argue that the issues of subjectivity, qualia and intentionality have become the main focusing points of the current debate. Neuroscientific research (e.g., direct excitation of the cortex, split-brain studies, blindsight and the timing of conscious experience) is taken to present empirical findings that are of direct relevance to philosophical theses. Yet, the authors are not overly optimistic. The results themselves are often in need of further explanation, and although there are converging lines of research, there is as of yet no agreement on what the study of consciousness actually amounts to. As they say, 'nobody is going to tell you the orthodox view, because there is no such thing' (p. 21).

Niiniluoto argues in the first chapter that scientific realism is a valuable position to consider because it has played an important role in rehabilitating the empirical study of human consciousness by holding, in contrast with positivism and instrumentalism, that science *can* disclose the true nature of unobservable mental events. Echoing a point made in the general introduction, Niiniluoto claims that given the current multitude of rival theories, science still cannot decide metaphysical controversies surrounding consciousness, for lack of a serious unified theory. To me, Niiniluoto's dismissal of instrumentalism is a first indication of an underlying theme that seems to pervade much of the book: Dennett's position on consciousness is a recurrent topic of attack.

Interestingly, in the second chapter, Dennett tries to defend himself against the accusation that his theory is a verificationist or eliminativist denial of consciousness. He claims not to be denying the reality of consciousness at all, but just to be pointing out that consciousness just isn't like what many people (philosophers, neuroscientists) think it is. The reality of consciousness consists in the 'cerebral celebrity' of distributed micro-processes in response to whatever is given as input. There is nothing over and above these micro-processes other than their leaving traces for further processing. Having effect, winning the cerebral competition, is all there is to becoming part of consciousness. No additional semi-mysterious properties (to be discovered by neuroscience) are required, as far as Dennett is concerned.

Patricia Smith Churchland and Vilayanur Ramachandran attack a central element of Dennett's position, namely his suggestion that in the case of 'filling in' phenomena, the brain just represents that there is more of the same. They provide many instances of filling in phenomena, related to blind spots and cortical and artificial scotomata, to show that instead of merely ignoring the gap, as Dennett would have it, the brain actually *provides* something. Dennett's claim is shown to be incorrect by cases where the filling in is quite contrary to the overall pattern of the presented image. In conclusion, Churchland

and Ramachandran argue that Dennett too readily accepts a thesis that is contrary to neuroscientific data because of his behaviorist background.

In chapter 4, another opponent of Dennett, Searle, provides his own view on the nature of consciousness, its relation to the brain and its main features needing explanation, which are, as he says he would like to think, 'just a form of common sense' (p. 101). Well-known by now is his view that the Turing test disposes one to make the two most common mistakes at the same time, namely to assume that consciousness can be analyzed behavioristically and computationally. Behaviorism is wrong because there is no necessary connection between inner states and behavior, so even if the behavior of a system is convincing, it need not be conscious. Computationalism cannot properly deal with consciousness because computation is an observer-relative notion. It depends on an observer who imposes a computational interpretation on some phenomenon. Since natural sciences deal with intrinsic features of the world (one of which is consciousness), computational properties are just not the right kind of concept with which to approach consciousness. Both 'mistakes' seem to be made by Dennett, but unfortunately Searle here refrains from discussing Dennett's views at all. At this point, given the aim of the book, one would like to see at least an indication of how empirical data could help to decide between two such completely opposing philosophical views. To his credit, Dennett does relate his views to models and data, thereby at least creating the possibility of being empirically refuted. Searle is content by calling his own point 'obvious' (p. 103).

The second part of the book examines some problematic properties of consciousness. Haugeland opens with a significant article on intentionality. He agrees with Searle that intentionality is intrinsic and subjective rather than observer-relative and objective à la Dennett. Yet, he tries to find some common ground by noting that both Dennett and Searle recognize the normativity of intentionality (by appealing to rationality (Dennett) or satisfaction conditions (Searle)). The question then is how a naturally evolved physical brain can be intrinsically normative. Haugeland argues that it is the commitment of a subject to the standards that operate within a domain that constitutes the intentionality of his or her states. This suggestion is, Haugeland claims, Dennettian in the sense that it involves taking a stance, though not from the outside, but by the subject itself. By taking a committed stance towards domain-regulating principles intentionality is constituted. Haugeland's suggestion is at the same time Searlian in that intentionality is seen to be intrinsic to the subject itself. Yet, Haugeland argues that animals do not have the intrinsic intentionality that Searle attributes to them since they do not submit themselves to any norms or constitutive principles on their own responsibility. In that respect, animals are on par with computers in having, as Haugeland calls it, ersatz intentionality.

There is, to be sure, a lot in this paper that asks for more explication. An important task, I think, would be to explain the ability to commit oneself without invoking normative notions. Yet, the interesting promise contained in Haugeland's article is that there is a way of characterizing intrinsic intentionality in terms of something else that is more susceptible to analysis.

Dennett, in chapter 6, concentrates on another problematic aspect of consciousness: qualia. He argues that the perception of color is a discriminative internal state of the brain that underlies various behavioral dispositions and that that is all. These internal discriminative states do not also have some special intrinsic properties. In short, though there sure seem to be qualia, there are not any genuinely qualitative properties. Dennett illustrates his suggestion by describing a color-detecting robot that can compare colors internally by relating representations *of* color, instead of using a colored representation. The way the robot would perform its task is not in any significant way different from the way we do it, Dennett claims.

Here we see Dennett making a point that was strongly attacked by Churchland and Ramachandran in chapter 3. Obviously, they granted, there are no colored representations in the brain. However, in the case of blindsight, the brain actually provides representations of a nature specific to the perception of visual phenomena, instead of merely representing *that* there is more of the same. Speaking about the special visual nature of certain representations does not entail that the brain produces actual images or colors to be perceived by an internal observer. But it does suggest that there is more going on than merely representing a color by a label such as 'green' or '11011'. What this 'more' actually amounts to, and how it is related to the emergence of a subjective experience, is quite unclear, but Dennett has not succeeded in showing that the attempt to find an *empirical* answer to such a question is essentially misguided.

Baars, opening the third part of the book dealing with models of consciousness, describes his own cognitive model in chapter 7. By comparing characteristics of conscious and non-conscious processes, Baars concludes that the basic question is how a serial, integrated and limited stream of consciousness can emerge from a nervous system that is mostly unconscious, distributed, parallel and of enormous capacity (p. 154). His answer is a global workspace theory. Specialist processes broadcast their results to the global workspace, and compete to access information displayed in it. He suggests that his model can explain many of the functional characteristics of consciousness (e.g., its relation to voluntary action, problem solving, etc.). Baars (p. 162) reports that Searle, in discussion, asked whether his model was truly about consciousness. To Baars the answer is yes, although he also claims that his aim is not to answer the Nagelian question of 'what it is like', but to understand the role of consciousness from a third-person perspective. A global workspace is necessary, though not sufficient, for consciousness. Searle's response is not reported, though it would be of considerable interest. If he found the answer convincing, then at least implicitly Searle would be acknowledging that computationally inspired models such as the global workspace theory might help to understand how consciousness arises in the brain. If he rejected the answer, then one might point out that Baars (in this chapter as well as in chapter 9) does attempt to specify neurobiological mechanisms that Searle himself supposes (p. 96) must be there in order for consciousness to arise.

Young, in chapter 8, goes into the neuropsychology of awareness in great detail. He makes clear at the outset that awareness comes in different kinds. After presenting data in the areas of visual field defects, achromatopsia, prosopagnosia, amnesia and anosognosia,

his general conclusion is that these deficits show that different aspects of awareness are lost in each specific case, and that there is no general perturbation of consciousness. Because of this, Young echoes a theme from Dennett in wondering "whether the subjective unity of conscious phenomena is not largely illusory" (p. 195). Indeed, Young warns against theories that treat consciousness as a single homogeneous entity. As Revonsuo notes in the introduction to this part of the book (p. 146), this puts Young in opposition to Baars' global workspace theory, but although one may speculate about discussions concerning this issue during the conference, there are no reports of it here.

The final part of the book addresses the future of consciousness research. Baars and Newman offer a neurobiological interpretation of the global workspace theory outlined in chapter 7. Especially, the suggestions concerning competition for access to the global workspace and the widespread dissemination of conscious information (broadcasting) can employ the expanding neurophysiological evidence. Neural structures associated with conscious wakefulness (the brainstem reticular formation and the nucleus reticularis) have an insufficient bandwidth to carry the information for a conscious full visual scene. Thus competition between potentially conscious perceptual contents for access to these neural structures seems likely. With respect to the broadcasting aspect of consciousness, Baars and Newman point, among other things, at the fact that the performance of novel tasks and the orienting response both require conscious attention and both involve widespread activity in the brain.

Tuomela addresses something of a side-issue with respect to the main theme of the book. He reviews the fate of folk psychology on the basis of a consideration of consciousness and argues for 'correctivism' instead of eliminativism. He discusses aspects of the traditional mind-body problem as related to self-awareness, qualia, intelligence and personality, and claims that one can be ontologically monist about these phenomena while continuing to use irreducible psychological predicates. He systematically examines twelve arguments for and against the elimination of folk psychology, which in itself is a welcome clarifying contribution to the eliminativist debate, and argues that none of these arguments is fully decisive alone, but together make the case for either eliminativism or full-blown anti-eliminativism rather weak. The option of correctivism remains alive and he attempts to strengthen it further by his 'thesis of correction': as long as human beings are viewed as thinking, feeling and acting persons, elimination is only possible to the extent that the topic is not changed. Though there may be a substantial change in the basic conceptual categories, the main topic of empirical investigation as given above need, by conceptual necessity, remain constant (p. 243).

Revonsuo deserves credit for his attempt, in the final chapter, to systematize and interrelate many of the issues and views regarding consciousness as discussed in the book. His starting point is the claim, made earlier in the introduction, that current theorizing about consciousness lacks a common core and he proposes a distinction between a philosophical and an empirical aspect of the problem of consciousness.

Philosophically, the problem is whether consciousness ultimately is a property of a neurobiological, functional, computational or completely different kind. Empirically, an

important issue is how the brain solves the binding problem. By focusing on the binding problem, Revonsuo is uniting with Young in asking how the phenomenal unity of consciousness is possible, given the divergence at the neurophysiological level.

Armed with this distinction, Revonsuo reviews five models of consciousness. Dennett's multiple drafts model gets criticized for being 'eccentric' in its answers to the ontological and binding problem. Consciousness is turned into a mere observer-relative label, whereas the binding problem is pushed away as originating in a mere seeming. In this way, Revonsuo concludes, the multiple drafts model distorts the meaning of consciousness and can therefore not provide the basis for a proper science of it. Dennett, of course, could simply repeat his defence, given in chapter 2, that he is being a realist about consciousness and is just stating that consciousness is not like what many people think it is. Yet, it seems to me, by treating the subjective experienced quality of qualia as mere seeming, Dennett *is* denying a central characteristic of the phenomenon under investigation. As long as empirical scientists see ways of investigating consciousness as traditionally conceived, Dennett's position is likely to be dismissed in the way of Revonsuo.

Jackendoff's computational model of consciousness leads, according to Revonsuo, to no satisfactory results. The underlying problem of all computational approaches, Revonsuo argues, is that computations play no causal or explanatory role in the system but is just a formal description of the system's operation. Agreeing with Searle that the notion of computation denotes nothing in nature itself but implies reference to an external interpreter, Revonsuo rejects the notion of computation as just not suited as a basis for theories of consciousness.

The remaining three models of consciousness Revonsuo deems to be more promising. The neuropsychological model is based on brain-injured patients having unconscious but not explicit knowledge of stimuli (among others, as discussed by Young in chapter 8). This model is taken by Revonsuo to be in agreement with the cognitive model, outlined by Baars in two earlier chapters, in assuming one common system that underlies all conscious phenomena. Finally, the neurobiological model (a label covering a variety of theories of, among others, Damasio, Crick, and Edelman) is of great interest because it pays close attention to the borderline between macro and micro levels of organization. Basically the binding problem is solved by means of the suggestion of a 'spotlight mechanism' based on time-locked neural activation or frequency-locked oscillations. A combination of these three models with a realist ontology is, according to Revonsuo, the way to go in the future study of consciousness.

This conclusion to the book may still be a long way from establishing the orthodoxy that, in the general introduction, was judged to be missing. Yet, all in all, one can safely say that the book presents a clear illustration of the current state of art in the interdisciplinary debate about consciousness. It reveals the stark contrast between certain positions, the intricacies of the discussion, the many loose ends and unsolved puzzles, but it also provides vivid instances of the growing inter-relation between philosophical views and empirical science.

