Introduction to Velmans, M. (ed.) (2018) Consciousness (Critical Concepts in Psychology) Volume 3: Cognitive and Neuropsychological Approaches to the Study of Consciousness Part 2, Major Works Series, London: Routledge

The Search for the Neural Correlates of Consciousness (NCC)

From the late 1950s onwards, Western psychological, neuroscientific and philosophical theories of mind and consciousness largely adopted one or another form of physicalist or functionalist reductionism—the view that consciousness is nothing more than a state or function of the brain. Francis Crick (1994), for example, made the point that in science, reductionism is both common and successful. Genes turned out to be nothing but DNA molecules, so in science, this is the best way to proceed. While he recognised that the experienced (first-person) "qualia" of consciousness pose a problem for reductionism, he suggested that in the fullness of time it may be possible to describe the *neural correlates* of such qualia. And, if we can understand the nature of the correlates we may come to understand the corresponding forms of consciousness. By these means, he claimed, science will show that "You're nothing but a pack of neurones!" (Ibid. p.3)

However, this view was overly simplistic. In science, correlates are not the same as causes, and neither correlation nor causation suffice to establish ontological identity. Given this, and the many other problems for reductionism posed by consciousness[1], contemporary neuroscientists have largely adopt a more cautious approach, which accepts that it might be possible to isolate the exact causal antecedents and neural correlates of a given conscious experience, without reducing that experience to either its antecedent causes or its correlates. The papers in this section introduce major phases and findings in this search.

The section begins with a review of seminal studies by the neurophysiologist Benjamin Libet (1982—R48). Subjectively, we seem to be immediately aware of what we attend to. However, experiments on the timing of conscious awareness by Benjamin Libet suggest that consciousness of input does not arise until at least 200 msec. after stimuli arrive at the cortical surface. Libet, et al. (1979), for example, found that direct microelectrode stimulation of the somatosensory cortex required a pulse train of at least 200 msec. duration before any conscious awareness of the stimulus was reported (pulse trains 10% shorter than this were not subjectively experienced). They also found that tactile stimuli applied to a finger were masked (prevented from entering consciousness) by microelectrode stimulation of the somatosensory cortex applied up to 200 msec. *after* the arrival of the tactile stimuli. On the grounds that one cannot prevent a stimulus from entering consciousness *after* it has done so, they concluded that at least 200 msec. of processing time are required to produce neural conditions adequate to support

---

[1] See Velmans (2009—R23) and the far more detailed analysis in Velmans (2009, chapters 1 to 5).

consciousness. The reason we do not experience any mismatch between experienced and actual stimulus arrival time appears to be that the brain records the actual time of arrival of the stimulus at the cortical surface. The brain then enters this into the representations of input time that it constructs—in spite of the fact that the representations themselves take about 200 msec. to construct—a form of "backward referral in time".

Francis Crick and Christof Koch (1990—R49) gave a useful summary of the discoveries in experimental psychology that have a bearing on the understanding of consciousness, but noted that the functional models provided by this form of investigation were not sufficiently precise to offer a unique description of the processes that support consciousness in the brain. They then proposed a series of questions that need to be answered to develop such an understanding along with a programme of research that might in principle begin to answer them. This paper and their subsequent collaborations had a major influence in fostering a new cognitive neuroscience of consciousness and with it the 'respectability' of *Consciousness Studies* within mainstream science.

Human conscious experience is normally well integrated. But how is this achieved? Physical objects and events are initially subject to a detailed (preconscious) analysis of physical features, and these features are coded in topographically distinct regions of the brain. Given this, the brain must somehow combine these features into representations that correspond to (and correlate with) the integrated objects and events that we experience. The contents of consciousness also change continuously, but they are bound together in time, appearing as a seamless, coherent flow in space and time. Singer (2013—R50) reviews evidence that the brain achieves such "binding" by the temporary association of neurons activated by currently present features into widely distributed, functionally coherent assemblies, which as a whole represent a particular content. Following this flexible strategy, a particular neuron can, at different times, participate in different assemblies just as a particular feature can be part of many different perceptual objects. This also allows for the rapid de novo representation of constellations that have never been experienced before because there are virtually no limits to the dynamic association of neurons in ever changing constellations. A currently favoured theory of the binding process suggests that this might be achieved by mutual entrainment of neuronal oscillations resulting in the synchronous or correlated firing of diverse neuron groups representing currently attended to objects or events, and, as Singer points out, the evidence for the existence of such binding processes, involving rhythmic frequencies in the 30 to 80 Hz region is now quite extensive.

Over the last 25 years or so, research on the neural markers of consciousness have become increasingly sophisticated, along with an understanding of the neural systems that normally support it in the human brain. Stanislas Dehaene and Jean-Pierre Changeux (2011—R51) survey a program of research that focuses on how an external or internal piece of preconscious information gains access to conscious processing, a transition characterized by the existence of a reportable subjective experience. They report that converging neuroimaging and neurophysiological data, acquired during minimal experimental contrasts between conscious and nonconscious processing, point to the following objective neural measures of conscious access: late amplification of relevant sensory activity, long-distance cortico-cortical synchronization at beta and gamma frequencies, and ‘‘ignition’’ of a large-scale prefronto-parietal network. They then go on to compare these findings to current theoretical models of conscious processing, including the Global Neuronal Workspace (GNW) model [2] according to which conscious access occurs when incoming information is made globally available to multiple brain systems through a network of neurons with long-range axons densely distributed in prefrontal, parieto-temporal, and cingulate cortices. The clinical implications of these results for general anesthesia, coma, vegetative state, and schizophrenia are also discussed.

Global Disorders of Consciousness

One of the obvious ways to determine the necessary and sufficient neural conditions for consciousness in the human brain is to isolate conditions that *abolish* it or globally *impair* it. Joseph Bogen (1995—R52) gives a detailed case in support of the Intralaminar Nuclei (ILN) of the thalamus having a critical role in the support of consciousness that draws on a wide range of resources, using evidence from anatomy, physiology, pathology, psychology and philosophy. For example, even small lesions in ILN can produce coma from which patients never recover. By contrast, cortical lesions, even as large as hemispherectomies, only abolish *some* contents of consciousness, not consciousness itself. While this does not establish consciousness to be located in ILN, it does suggest that these serve as important gateways to it in the normally functioning human brain.

Giacino, Fins**,** Laureys and Schiff (2014)[3] provide a very thorough survey of global disorders of consciousness after acquired brain injury, such as coma, the vegetative state, minimally conscious states such as akinetic mutism, acute confusional states, and locked-in syndrome. They review the consciousness-altering pathophysiological mechanisms involved, the clinical syndromes they present, and discuss novel diagnostic and prognostic applications of advanced neuroimaging and electrophysiological procedures. Broadly speaking, disorders of consciousness can arise from both focal injuries that induce widespread functional changes and from more global injuries, from both direct perturbations of neural systems that regulate arousal and awareness and indirectly from disruptions in the connections between these systems—and they go on to specify the details of such injuries and consequent changes. They conclude with a provocative discussion of bioethical and medico-legal issues that are unique to this population and have a profound impact on care, as well as raising questions of broad societal interest.

---

[2] The Global Neuronal Workspace (GNW) model is a development of the Global Workspace (GW) model originally advocated by Baars (1988)—see R38. Dehaene (2014) also provides an excellent, detailed account of this program of research.

[3] Unfortunately, the reprint cost of this paper was too expensive to include in this collection, but it is available online at http://www.nature.com/nrneurol/journal/v10/n2/full/nrneurol.2013.279.html

Some challenges for NCC research

Most contemporary theories of consciousness in humans give prominence to the role of the thalamo-cortical complex. However, in his *Behavioral and Brain Sciences* target article "Consciousness without a cerebral cortex" Bjorn Merker (2007—R53) challenges the necessity of a cerebral cortex. Drawing on evidence about the functional organization of vertebrate brains ranging from comparative neurology to experimental psychology and from neurophysiology to clinical data in humans he presents a detailed case for the central role of the *upper brain stem system* in the support of consciousness and the control of behavior, a view pioneered by Wilder Penfield (reviewed in Penfield, 1975). This system, which extends from the roof of the midbrain to the basal diencephalon, integrates the massively parallel and distributed information capacity of the cerebral hemispheres into the limited-capacity, sequential mode of operation required for coherent behaviour. It maintains special connective relations with cortical regions that are involved in attentional and conscious functions, but is not rendered nonfunctional in the absence of cortical input. As he notes, this helps to explain the purposive, goal-directed behaviour exhibited by mammals after experimental decortication, as well as the evidence that children born without a cortex are conscious. On the basis of such evidence, he concludes that this upper brain stem system retained a key role throughout the evolutionary process by which an expanding forebrain—culminating in the cerebral cortex of mammals—came to serve as a medium for the *elaboration* of conscious contents but not for its *existence*.

Such caveats highlight the importance of positioning human consciousness within a broader evolutionary context, with consequences for our understanding of consciousness in non-human animals (see also R47, R48). More fundamentally, it also points to the importance of distinguishing the necessary and sufficient conditions for the *existence* of consciousness (of any kind) from the *added* conditions required to support the many *forms* it can take—which are easily confounded in the search for NCC (c.f. Velmans, 2012).

In another challenging paper, Aru et al. (2012—R54) give a thoughtful review of the problems of separating the precise neural correlates of a given conscious content from its neural *prerequisites* (NCC-pr) and neural *consequences* (NCC-co), which can be confounded if one simply compares perceptual conditions in which a given content is present with similar conditions in which it is absent (the commonly used method of "contrastive analysis"). They review evidence of such confounds in the literature and propose experimental strategies to untangle them. They suggest, for example, that the most straightforward way to disentangle NCC-pr from NCC is to directly manipulate the NCC-pr processes, and compare the neural signatures that are common to all of them. Given that the NCC-co are after-effects of conscious perception, it is logically possible to observe NCC even when NCC-co are not elicited. Such dissociations might not occur very often in the normally functioning brain, however lesion studies might reveal them, or in some situations one might be able to repress NCC-co experimentally with transcranial magnetic stimulation (TMS)—and they go on to give examples of the data that such studies provide.

A similarly cautious message is conveyed by the extensive, recent review of progress and problems of NCC research by Koch, Massimini, Boly, and Tononi (2016).[4] Koch et al. distinguish between content-specific NCC, full NCC, and background supporting conditions for consciousness. Content-specific NCC are the neural mechanisms whose activity determines a particular phenomenal distinction within an experience. The full NCC are the neural substrates supporting conscious experiences in their entirety, irrespective of their specific contents. The brainstem reticular formation, paramedian thalamus and perhaps parts of the postero-medial cortex are likely to provide the *background* conditions for full consciousness. They also note that at least part of the neural activity that co-varies with the perception of a particular conscious content reflects processes that precede or follow the experience, such as selective attention, expectation, self-monitoring, unconscious stimulus processing, task planning and reporting, rather than the experience itself (see R54). Once these functional conditions are controlled for the reliability of most neural markers currently thought to be indices of consciousness weakens. For example, many studies of NCC did not carefully dissociate conscious visibility from selective attention (see also R35). When this is done, the high-range gamma synchronization (often thought of as a marker of consciousness) correlates with attention, independent of whether the participant saw the stimulus, whereas mid-range gamma synchronization correlates with the visibility of the stimulus. Even worse, gamma synchrony can persist or even increase during early Non-REM sleep, during anaesthesia, or during seizures. They also survey evidence, that overall, there is only a weak correlation between consciousness and thalamic activity, making this an unreliable marker of consciousness. In the light of this, they go on make various suggestions about how NCC research should proceed, argue for a few alternative candidate NCC (a posterior cortical activity "hot-spot", and various measures of "cortical integration"), and conclude that the low-voltage fast activity observed with EEG recordings performed during attentive wakefulness, also known as activated or desynchronized EEG—which was one of the first candidate electrophysiological indices of consciousness—is still one of the most sensitive and useful markers available.

The Divided Brain

The left and right hemispheres of the brain are connected by the cerebal commissures, and research with cerebral commissurotomy patients (where the commissures have been sectioned) proved to be very useful in determining the respective functions of the two halves of the brain. It also raised some fundamental theoretical/philosophical questions: Could consciousness itself be divided by this operation? And, if so, would such patients have a distinct left-brain and right-brain consciousness? "Split brain" patients were all advanced epileptics in whom an extensive midline section of the cerebral commissures had been carried out in an effort to contain severe epileptic convulsions not controlled by medication. Studies of such patients, pioneered by Roger Sperry and his colleagues Philip Vogel, Joseph Bogen and Michael Gazzaniga exploited the fact that the left half of the visual field projects onto the right

[4] This major review has many, important implications for NCC research. Unfortunately, the reprint cost of this paper was too expensive to include in this collection, but it is available online at http://www.nature.com/nrn/journal/v17/n5/abs/nrn.2016.22.html

hemisphere while the right half of the visual field projects onto the left hemisphere; the left hemisphere also controls the right hand, while the right hemisphere controls the left hand. This allows information to be projected to only one hemisphere, and a way of measuring its response—a way to assess the functioning and associated experience of each hemisphere. Sperry (1968—R55), and Sperry, Gazzaniga & Bogen (1969—R56) provides good, initial reviews of this research along with descriptions of which functions are divided and which are spared by this operation.

The most remarkable effect of sectioning the neocortical commissures in this way is the apparent *lack of effect* on normal behaviour. Nevertheless, in these carefully controlled studies, patients behaved in many ways as if they had two independent streams of conscious awareness, one in each hemisphere, each of which was cut off from the experiences of the other. Each hemisphere seemed to have its own separate, private sensations, perceptions, concepts, and its own impulses to act, with related volitional, cognitive, learning experiences, and memories that were inaccessible to the recall processes of the other. The two hemispheres also turn out to be specialised to function in somewhat different (but overlapping) ways. For example in right handed patients, the left hemisphere typically controls speech and has superior linguistic functioning, while the right hemisphere has superior control of visual-spatial operations. The studies explored these and many other fine grained differences in how the two halves of the brain differ and how they interact, which subsequently fed in to an extensive, on-going research program into left brain versus right brain differences.

The recent review chapter by Colvin, et al (2017—R57) brings developments in this research up-to-date, with an added focus on why it might that, in spite of evidence for a distinct left and right brain consciousness, comissurotomy patients do not appear to exhibit split consciousness in their normal behaviour. Following Gazzaniga's interpretation of the evidence, they argue that the left hemisphere, which normally controls speech and the ability to report, simply responds on the basis of all the information available to it—and, lacking access to right hemisphere experience, has no sense that it's consciousness is split.

The Reintroduction of First-Person Methods

As Boring (1953—R14) observed, introspective methods never really left psychology, even during its behaviourist period. However with the re-emergence of Consciousness Studies in the 1990s and the development of more sophisticated neuroimaging methods (PET, fMRI etc.), there was a renewed interest in developing more refined first-person (introspective) methods for the investigation of conscious experience that could, in principle, complement the finer detail in which neuroscientific methods were beginning to reveal the dynamics of the brain. This renewed interest was marked for example by the launch of a new interdisciplinary journal, *Phenomenology and the Cognitive Sciences* In 2002, and new collections of readings and major reviews of first-person methods including Varela & Shear (1999), Velmans (2000b), Petitmengin (2009), Hurlburt (2011), and Price and Barrell (2012—R61, 62, 63).

Neurophenomenology

In a review of the state-of-play in the mid 1990s, along with suggestions on how to move things forward, Francisco Varela (1996—R58) drew a useful map of where 16 authors with extensive writings on how to understand consciousness positioned themselves on two dimensions: (a) stress on first-person accounts of phenomenology being essential, versus neo-reductionism or eliminativism (first-person accounts having no place in science at all); and (b) in terms of their commitment to the empirical solubility of the "hard problem" of consciousness, ranging from functionalist reductionists who believe all problems will be solved by discovery of the processes (cognitively or physiologically specified) with which consciousness is most closely associated, versus mysterians who argue that a full understanding of consciousness is beyond human mental capacities. Varela positioned himself as (a) defending the importance of phenomenology and (b) not committed to either functionalist reductionism or mysterianism. Instead, he advocated *neurophenomenology*, a suggested methodological route through the "hard problem" in which refined forms of phenomenological investigation (developed within the tradition of European phenomenology) could be related to the more detailed accounts of neurodynamic activities accompanying such phenomenology that were becoming available from investigations of the brain (see also R61). At the *Centre National de Recherche Scientifique* in Paris, he also initiated a research program to develop such first-person methods and research, and specified some of the characteristics of that program in this paper.

Antoine Lutz and Evan Thompson (2003—R59) provide a nice illustration of how this program can be carried out in practice, in a way that synergises, (1) first-person data from the careful examination of experience with specific first-person methods, (2) formal models and analytical tools from dynamical systems theory, grounded on an embodied-enactive approach to cognition, and (3) neurophysiological data from measurements of large-scale, integrative processes in the brain. They also elaborate on different forms of consciousness (viewed from a phenomenological perspective), the skills required to develop an attentive awareness of one's own experience (similar to mindfulness-awareness in the Buddhist tradition), and (second-person) methods whereby descriptions of such experiences can be shared in experimental investigations. In their selection of salient neural indices of cognitive-phenomenal (conscious) states, they advocate the relevance of collective variables that describe the emergence and change of patterns of large-scale neural integration—but stress that such states must be understood as enactive, embodied, and embedded in the environment. They then describe a neurophenomenological pilot study that incorporates these principles.

The most extensive research program in this tradition has been carried out by Claire Petitmengin and her colleagues. Bitbol and Petitmengin (2017—R60) review this program and its broader implications. They point out that, for Varela, the phenomenological roots of neurophenomenology provide more than an experimental method. Given its stress on the epistemic primacy of lived, embodied experience, it can be seen as a way of bypassing conventional first-person ('subjective') versus third-person ('objective') ways of construing the mind/body problem. Instead, mind and body are viewed as two ways of ordering and selecting aspects of a single flux of lived experience, which lead, they argue, to a form of *deep neurophenomenology*, a potentially transformative process, in which conventional mind/body problems simply do not arise.[5] They document the detailed techniques of the micro-phenomenological interview (previously called the "elicitation interview"), which enable fine-grained recollection and description of specific aspects of one's own lived experience, allowing identification of its generic structures and its micro-dynamic changes, which usually remain unnoticed and unrecognized. The method also includes a set of objective verbal and non-verbal criteria for checking its own effectiveness. They give examples of the diverse situations in which the procedure has allowed the detection of previously unnoticed, but consistent micro-dynamics in diverse experiences, ranging from studies of creative thinking where the method revealed the progressive transformation of a fuzzy and blurred feeling into a "clear and distinct idea", to clinical applications, where epileptic patients learnt to become aware of early signs announcing the arrival of an epileptic seizure—providing a key to new cognitive therapies. They go on to give a critique of the various ways that the potential of introspective methods have been misunderstood, and give a detailed analysis of how the fine-grained nature of the micro-phenomenological interview may be an appropriate accompaniment to the increasingly fine-grained study of the neurophysiological dynamics of the brain.

The use of complementary first- and third-person perspectives

The view that consciousness and its neural accompaniments are best thought of as two manifestations of one underlying, psychophysical process that can be known in two complementary, first- and third-person ways has roots in the philosophical writings of Theodore Fechner (see R4) and older roots in the dual-aspect monism of Spinoza (1677). In the modern literature, it finds its most complete, experimental expression in the work of Donald Price and James Barrell (2012), which introduces and reviews a form of *experiential neuroscience*, a program of research that reintegrates studies of experience with psychology and neuroscience in which they have engaged for over 40 years.

In their chapter on "Developing a science of human meaning and consciousness", Price and Barrell (R61) introduce the philosophical and methodological origins of their approach, pointing out that both first- and third-person data have their source in how we *experience* the phenomena of the world as well as our innermost thoughts, feelings and meanings—a self-evident truth that is largely ignored by Western materialist philosophy and science. They briefly review other research programs with similar aims (e.g. introspection, phenomenology, and Hurlburt's *Descriptive Experience Sampling* technique) and introduce two important, added ingredients

---

[5] The epistemic primacy of *experience* in knowledge and consequently its role in the *creation* of (and potential resolution of) the subjective versus objective distinction was also advocated by the *neutral monists* (see James, 1902—R5), as well as later readings in this section, although these alternative approaches offer different ontologies.

required for a fully integrated first- and third-person approach: quantitative psychophysics, and the experiential exploration of not just the sensed qualities of experience, but also their personal *meanings* which are essential not just to an understanding of the experiences themselves, but also to consequent behavior and even psychophysiological response. They go on to give details of their experiential paradigm, which applies structured methods for (1) attending to the content of specific experiences, (2) reporting that content, (3) understanding that content through the discovery of common factors, and (4) applying quantitative methods, drawn largely from psychometrics and psychophysics, to test the generality and functional relationships between common factors.[6] They go on to discuss the relevance of such studies to a more complete understanding of the associated neuroscience, pointing out the limitations of reductionist approaches, and elaborating on ways in which neuroscience and experiential science can enhance each other, for example within a reflexive model of these relationships (Velmans, 2009) which accepts that neither first- nor third-person accounts are observer-free or incorrigible, that a complete account of the operations of mind require both—and they go on to describe the potential of such an approach for a broader understanding of the human condition.

In their chapter on 'Human pain and suffering' (R62) Price and Barrell give a very detailed demonstration of how their program works in practice. As they note, pain is an *experience* that has both sensory and affective dimensions. Not merely unpleasant, it can be the kind of unpleasantness that makes people scream, fight, undergo crippling, disfiguring operations, or commit suicide. These intense, complex dimensions need to be made explicit, which a well-formed experiential analysis can provide. Once explicit, such dimensions can provide a sound basis for investigations of pain-related neural pathways, neurotransmitters, and integrative centers of the brain. For example, studies by Price and his colleagues in the 1980s and 1990s of many different forms of pain including musculoskeletal pain, neuropathic pain, cancer pain, and labour pain isolated three common factors: (1) unique sensory qualities (to a given form of pain) that are like those which occur during tissue damaging stimulation, (2) a closely related meaning of intrusion and/or threat, and (3) a related feeling of unpleasantness and/or other possible negative emotion(s). Not surprisingly, factors (2) and (3) can vary independently of (1). Hot spicy foods for example are experienced as intrusive or damaging by some individuals but as pleasant by others. The meaning attributed to a given pain (2) also influences how unpleasant it is (3). Price, Harkings and Baker (1987) for example found that in a study of labour pain, the unpleasantness of pain was heavily influenced by whether women were focused mainly on avoiding the pain or focused on having the baby, with the latter rating pain unpleasantness at approximately one-half of those focused on the pain or on avoiding it and that this

---

[6] For example, once factors central to a given experience are found they can be scaled into measurable dimensions and then manipulated as dependent or independent variables in conventional experimental designs to explore their functional relationships. Causal relationships can also be explored using experimental designs that use structural equation modeling to isolate which factors are necessary and sufficient for the existence of other factors, along with other techniques that can track causal dependencies within a given experiment.

was true at every stage of labour. In short cognitive-evaluative factors also influence pain unpleasantness, a finding that has clear therapeutic implications.

Once elaborated in this way, it is also easier to distinguish the pain sensation as such from consequent suffering. The immediateness and unpleasantness of pain is also often accompanied by a *desire* to reduce or terminate it, and a level of *expectation* that it can be reduced, and these dimensions of experience turn out to be causal factors in pain-related emotional feelings such as depression, anxiety, frustration and fear, along with emotions associated with the interruption of one's life activities, the difficulty of enduring the pain over long periods, and concern for future consequences. The functional relationships among these complex factors can be clarified a) by studies that examine the successive psychological stages of pain processing, e.g. using a form of path analysis (structural equation modeling), and b) studies that demonstrate the selective effects of personality traits and demographic factors on suffering (as opposed to the immediate pain experience). To complete the circle of understanding Price and Barrell then give a *very detailed analysis* of how these psychological functional dependencies can be related to a complex, testable, neural model of pain-related processing (based on Price's own extensive neuropsychological research).

In their chapter on 'Second pain: a model for explaining a conscious experience?' (R63) Price and Barrell point out that second pain has unique features not predictable by either the nature of physical stimuli or the responses of peripheral pain receptors (connected to C axons). A single intense stimulus often gives rise to impulse input from fast conducting A-delta axons followed by impulses from slow conducting C axons, which respectively elicit first and second pain. Unlike sharp or pricking first pain, second pain has burning, aching or "dull" qualities, exhibits slow temporal summation, becoming stronger with repeated stimuli, often outlasts a stimulus that evokes it, and spreads into body regions outside the area of the stimulus as the pain becomes more intense. These characteristics are not easily explained by the nature of the physical stimulus alone, and require studies of neurons, pathways and interactions within the central nervous system. Price and Barrell then provide a very detailed account of the neural mechanisms involved—and provide evidence that neural activity in widely distributed areas within the spinal cord and brain are the causes and correlates of these different pain characteristics, with distinct activity patterns that closely correspond to each of the characteristics identified by first-person methods. They conclude with a discussion of the wider philosophical and methodological implications of this research—and overall provide what is perhaps the best example of how to combine an "outer psychophysics" with an "inner psychophysics" in the manner envisaged by Fechner (R8) and advocated by Wundt (R9).

The last reading in this section on 'What consciousness does' (Velmans, 2009, ch. 13—R64) demonstrates how complementary first- and third person perspectives associated with dual-aspect monism can provide a plausible account of how to make sense of one of the hardest, most enduring problems of consciousness—the causal interactions of consciousness and brain.[7] As he notes, conventional medicine takes the physical causation of physical effects for granted, psychiatry takes the physical causation of mental effects for granted, many forms of psychotherapy take the causal effects of mental states on other mental states for granted, and psychosomatic medicine, psychoneuroimmunology etc., assume that mental states can have causal effects on physical states. Given the extensive evidence for all these forms of causation, how are we to make sense of them? The chapter reviews the main problems with dualist and materialist reductionist accounts, and develops the case for a non-reductive, dual-aspect monist alternative, in which first-person (experiential) and third-person (neurophysiological) observable features of the operations of mind (and consequent first- versus third-person accounts) are complementary views of one, underlying, psychophysical form of human information processing (roughly analogous to the way that electricity and magnetism may be thought of as complementary features of electromagnetism). This provides a simple way of making sense of all four forms of physical/mental causation. Operations of mind viewed from a purely third-person perspective appear as forms of physical→physical causation, operations of mind viewed from a purely first-person perspective appear as forms of mental→mental causation, and cases of mind/body interaction (physical→mental and mental→physical) can be seen as *mixed-perspective accounts* involving *perspectival switching.* All of these can be understood as different views (or a mix of views) of a single, psychophysical form of information processing developing over time. In providing a common psychophysical ground for brain operations and associated experiences, such a process also provides the "missing link" required to explain psychosomatic effects. The chapter then goes on to demonstrate how this approach can resolve the "causal paradox"—that, viewed from a purely third-person perspective, consciousness appears causally ineffective (epiphenomenal) while viewed from a first-person perspective consciousness and its operations appear to be crucially important to much of human life. [8]

Free Will

Whether humans have free will or whether their actions are entirely determined by the laws of physics, physiology, and the accidents of heredity and environment has long been debated within philosophy and science. In modern consciousness studies however the debate was energized by the findings of Benjamin Libet on how feelings of wishing or deciding to act relate to prior processing in the brain. Much of this work was summarised in Libet (1985—R65) a target article in the *Behavioral and Brain Sciences*. Libet demonstrated that a slow negative shift in electrical potential (known as the "readiness potential"), which was known to signal preparation for action, preceded the experienced wish to act by around 350 milliseconds, suggesting that the brain was preparing to act, before the conscious experience even arises! Given the theoretical implications of this as well as ethical and legal implications, this finding was subject to extensive critique, debate and consequent research.

---

[7] Originally developed in Velmans (2000a) this analysis also formed the basis of Velmans (2002a) a target article for a special issue of the *Journal of Consciousness Studies*, followed by 12 commentaries and two responses (Velmans, 2002b, 2003).

[8] Although arrived at independently, the analysis is intriguingly similar to the dual-aspect monism developed by Romanes (1885—R21) to counter Huxley's epiphenomenalism (R20).

Chris Frith (2013—R66) provides an insightful review of this research, arguing that many of the problems of free will arise from mixing up two perspectives from which volition can be viewed (see also R21, R61, R64). From the third-person view, volitional behaviour is internally generated, rather than being determined by the immediate environmental context, and is therefore, to some extent, unpredictable. From the first-person view, our experience of volitional behaviour includes a vivid sense of agency. We feel that, through our intentions, we can cause things to happen and we can choose between different actions. However, our experience of agency is not direct. It depends on sub-personal inferences derived from prior expectations and sensations associated with movement. As a result, our experiences and intuitions about volition can be unreliable and uncertain. Nevertheless (viewed from a first person perspective) our experience of agency is not a mere epiphenomenon. For example, anticipation of the regret we might feel after making the wrong choice can alter behaviour—and the strong sense of responsibility, associated with agency, has a critical role in creating social cohesion and group benefits. In arguing for these views, Frith examines the relationship between free choice, determinism, flexibility, and unpredictability, the nature of endogenous self-generated behaviour, and the extent to which one can predict endogenously generated behaviour on the basis of prior brain activity. He then turns to research on how volition is experienced, which suggests that, when all is going well, we are simply aware of our intention to act and of the effect of our action on the world (the outcome) as a check that our intention has been fulfilled. The contingent relation between the experienced intention and the outcome has also been shown to contribute in various ways to our experience of *agency*. Frith suggests that it is our experience of agency that gives us our sense of *responsibility*, which, in turn, can be responsible for feelings of regret, along with cultural phenomena such as praise and punishment. He concludes by discussing the wider effects of culture on the experience of agency and the linked differences between human experiences of agency and that of other animals. [9]

Although much of the free-will research was energized by Libet's experiments on the lag between neural preparation to act and the experienced wish to act, in the last five years or so, the understanding of Libet's research has changed. Schurger (2017—R67) brings this topic (including a review of neurological disorders of volition) up to date. In particular, he reviews evidence that the complex brain activities preceding and contributing to an experienced decision to act (signaled by the readiness potential and other markers) continue to develop over time, and that the final neural decision to act in a chosen way may be much closer to the experience to do so than previously believed. That said, he also notes that, in recent years, debates about whether will is "free" no longer depend on whether will is "conscious", both in psychology and philosophy of mind (see also R64 where the implications of this for an understanding of conscious causation are discussed). He concludes that the processes that produce a *conscious feeling* of wanting to move a body part are independent from the

---

[9] Haggard (2008) also provides an excellent review along with a discussion of how voluntary action relates to consciousness, and some interesting comments about what this implies for social and legal responsibility. Unfortunately, the reprint cost of this paper was too expensive to include in this collection, but it is available online at http://www.nature.com/nrn/journal/v9/n12/full/nrn2497.html

mechanisms involved in actually bringing that movement about. However, these different processes might be very tightly coupled, and the feeling of intending might emerge at the same time that the movement is being triggered, at least under normal everyday circumstances. Overall, this leads to a form of compatibilism, in which first-person feelings (and accounts of those feelings) are complementary to third-person observations (and accounts of those observations) of associated brain activities. This, in turn, opens up the possibility that feelings of having freedom to act (within existing constraints) are compatible with the operations of the underlying processes that govern those choices—a route through the free-will versus determinism debate within philosophy of mind (Velmans, 2003).

VOLUME 3 TABLE OF CONTENTS

The Search for the Neural Correlates of Consciousness (NCC)

awareness', *American Psychologist*, 23, 723–733.

R56. Roger W. Sperry, Michael, S. Gazzaniga, and Joseph, E. Bogen (1969) 'Interhemispherical relationships: the neocortical commissures; syndromes of hemispherical disconnection', in P.I. Vinken and G.W. Bruvn (eds.), *Handbook of Clinical Neurology*, Amsterdam: North-Holland Publishing Company, pp.273-290.

R57. Mary K. Colvin, Nicole, L. Marinsek, Michael B. Miller, and Michael S. Gazzaniga (2017) 'Split-brain cases', in S. Schneider and M. Velmans (eds.), *The Blackwell Companion to Consciousness Second Edition*, Chichester, West Sussex: Wiley-Blackwell, 634-647.

The Reintroduction of First-Person Methods

Neurophenomenology

R58. Francisco J. Varela (1996) 'Neurophenomenology: a methodological remedy for the hard problem', *Journal of Consciousness Studies*, 3, pp. 330–350.

R59. Antoine Lutz and Evan Thompson (2003) 'Neurophenomenology: Integrating subjective experience and brain dynamics in the neuroscience of consciousness', *Journal of Consciousness Studies*, 10, 9–10, pp. 31–52.

R60. Michel Bitbol and Claire Petimentingen (2017) 'Neurophenomenology and the micro-phenomenological interview', in S. Schneider and M. Velmans (eds.), *The Blackwell Companion to Consciousness Second Edition*, Chichester: Wiley-Blackwell, pp. 726-739.

The use of complementary first- and third-person perspectives

R61. Donald D. Price and James, J. Barrell (2012) 'Developing a science of human meaning and consciousness', in D. D. Price and J. J. Barrell, *Inner Experience and Neuroscience: Merging Both Perspectives*, Cambridge, MA: The MIT Press, pp. 1-30.

R62. Donald D. Price and James, J. Barrell (2012) 'Human pain and suffering', in D. D. Price and J. J. Barrell, *Inner Experience and Neuroscience: Merging Both Perspectives*, Cambridge, MA: The MIT Press, pp. 150-193.

R63. Donald D. Price and James, J. Barrell (2012) 'Second pain: a model for explaining a conscious experience?' in D. D. Price and J. J. Barrell, *Inner Experience and Neuroscience: Merging Both Perspectives*, Cambridge, MA: The MIT Press, pp. 195-207.

R64. Max Velmans (2009) 'What consciousness does', in M. Velmans, *Understanding Consciousness Second Edition*, London: Routledge, pp. 300-326.

Free will

R65. Benjamin Libet (1985) ‘Unconscious cerebral initiative and the role of conscious will in voluntary action’, *Behavioral and Brain Sciences*, 8, 529-539.

R66. Chris Frith (2013) ‘The psychology of volition’, *Experimental Brain Research*, 229, 3, 289-299.
R67. Aaron Schurger (2017) ‘The neuropsychology of conscious volition’, in in S. Schneider and M. Velmans (eds.), *The Blackwell Companion to Consciousness Second Edition*, Chichester: Wiley-Blackwell, pp. 695-710.

Additional References (not already included in the Table of Contents)

Crick, F. (1994) *The Astonishing Hypothesis: The scientific search for the soul,* London: Simon & Schuster.

Dehaene, S. (2014) *Consciousness and the Brain: How the Brain Codes our Thoughts,* New York: Viking.

Giacino, J. T., Fins, J. J.**,** Laureys, S. and Schiff, N. D. (2014) Disorders of consciousness after acquired brain injury: the state of the science, *Nature Reviews Neurology*, 2014, 10, pp. 99-114.

Haggard, P. (2008) Human volition, *Nature Reviews: Neuroscience*, 9, 934-946.

Hurlburt, R. T. (2011) *Investigating Pristine Inner Experience: The Descriptive Experience Sampling Method*, Amsterdam: John Benjamins.

Koch, C., Massimini, M., Boly, M. and Tononi, G. (2016) Neural correlates of consciousness: progress and problems, *Nature Reviews Neuroscience*, 17, 307-321.

Libet, B., Wright Jr., E.W., Feinstein, B. and Pearl, D.K. (1979) Subjective referral of the timing for a conscious experience: A functional role for the somatosensory specific projection system in man, *Brain* 102, 193-224.

Penfield, W. (1975) *The Mystery of the Mind: A critical study of consciousness and the human brain*, Princeton: Princeton University Press.

Petitmengin, C. (ed.) (2009) *Ten Years of Viewing from Within: The Legacy of F. J. Varela*, Journal of Consciousness Studies Special Issue, Vol. 16 (10-12). Exeter, UK: Imprint Academic.

Price, D. D., Harkins, S. W., and Baker, C. (1987). Sensory-affective relationships among different types of clinical and experimental pain, *Pain,* 28**,** 291-299.

Spinoza, B. (1677[1876]) *The Ethics*, Reprinted in *The Ethics of Benedict Spinoza*, New York: Van Nostrand.

Varela, F. and Shear, J. (1999) *First-person Approaches to the Study of Consciousness*. Exeter, UK: Imprint Academic.

Velmans, M. (2000a) *Understanding Consciousness*: London: Routledge.

Velmans, M. (ed.) (2000b) *Investigating Phenomenal Consciousness: New Methodologies and Maps*, Amsterdam: John Benjamins.

Velmans, M. (2002a) How could conscious experiences affect brains? *Journal of Consciousness Studies*, **9**(11), 3-29.

Velmans, M. (2002b) Making sense of causal interactions between consciousness and brain, *Journal of Consciousness Studies*, **9**(11), 69-95.

Velmans, M. (2003) Preconscious free will, *Journal of Consciousness Studies*, 10(12), 42-61.

Velmans, M. (2009) Understanding Consciousness Second Edition, Routledge/Psychology Press, Hove: East Sussex.

Velmans, M. (2012) The evolution of consciousness. In D. Canter and D. Tunbull (eds.) *Contemporary Social Science. Special Issue on Biologising the Social Sciences: Challenging Darwinian and Neuroscience Explanations*, 7(2), 117-138.