1

A Brief History of the Scientific Approach to the Study of Consciousness

Chris D. Frith and Geraint Rees

COPYRIGHTED MATERIAL

The Origin of Consciousness Studies: René Descartes

The attempt to develop a systematic approach to the study of consciousness begins with René Descartes (1596–1650) and his ideas still have a major influence today. He is best known for the sharp distinction he made between the physical and the mental (Cartesian dualism). According to Descartes, the body is one sort of substance and the mind another because each can be conceived in terms of totally distinct attributes. The body (matter) is characterized by spatial extension and motion, while the mind is characterized by thought. This characterization of the mind also renders it private, a precursor of the distinction between the first-person and the third-person perspectives. Today, most scientists do not accept dualism, instead believing that mind somehow emerges from the physical properties of the brain. However, the distinction between mind and matter is still perceived as being so clear-cut that explaining how mind can emerge from matter, and reconciling the first-person and third-person perspectives, remain the hardest problems facing the student of consciousness.

Some consider that Descartes has impeded the scientific study of consciousness, since his development of dualism placed consciousness outside the domain of science. However, Descartes was an interactive dualist and, as such, was the first to think seriously about the neural correlates of consciousness. He recognized that the brain has a key role for sensory input and motor output, but this did not make it the basis of mind. He considered that non-human animals did not have minds, but were unthinking automata for which a brain was sufficient. There is an interesting parallel here with current distinctions between conscious and unconscious processes. For Descartes, consciousness was a state of mind, with the brain having a role restricted to nonconscious processes. Nevertheless, the brain had a key role in linking matter and mind. Physical bodies in the world have an impact on the sense organs. This impact creates motion in the body's nervous system that is somehow translated into the mind's experience of color, sound, and other sensations. These motions are transmitted to the pineal gland where they act as cues to the rational soul, enabling this to have specific types of conscious experience or ideas. We now know that Descartes was wrong about the importance of the pineal gland. But his account is not that different from recent proposals that, for example, neural activity in the fusiform region of the brain somehow leads to the conscious experience of a face.

The Blackwell Companion to Consciousness, Second Edition. Edited by Susan Schneider and Max Velmans.
© 2017 John Wiley & Sons Ltd. Published 2017 by John Wiley & Sons Ltd.

Descartes also made a distinction between what would now be called “bottom-up” and “top-down” processes. The passions, such as joy and anger, agitate and disturb the mind. Conflicts between the passions and the will occur when the body (bottom-up) and the soul (top-down) cause opposing movements in the pineal gland, that unique structure in the brain where mind and body interact. The interplay between top-down and bottom-up processes in determining the outcome of cognitive processes remains a common motif in contemporary cognitive neuroscience.

After Descartes

Since Descartes much effort was devoted in trying to put the physical and the mental back together again. Baruch Spinoza (1632–77) proposed that the mental and the physical are different aspects of the same substance (*dual aspect theory*), while Gottfried Leibniz (1646–1716) proposed that the mind and the body were separate substances, but constructed from the outset to run together in perfect harmony (*psychophysical parallelism*). George Berkeley (1685–1753) denied the possibility of mindless material substances (*immaterialism*). He proposed that things could only exist through being a mind or through being perceived by a mind. In contrast *materialism* holds that matter is fundamental and is the cause of mental events. This is an ancient idea championed by, among others, Julien Offray de la Mettrie (1709–51) in his book *L'homme machine*. La Mettrie extended Descartes's idea of animals as automata to man. In particular, he proposed that conscious and voluntary processes result simply from more complex mechanisms than involuntary and instinctive processes. This is, in essence, the belief held by many of us who are searching for the neural correlates of consciousness in the twenty-first century.

John Locke (1632–1704) and the empiricist philosophers who followed him were less concerned with the mind–body distinction and more concerned with the problem of knowledge: how the mind learns about the world. Locke contrasted *outer sense*, the mind's experience of things, with *inner sense*, the mind's reflective experience of its own experience of things. He also recognized the importance of the *association of ideas*, a concept taken further by David Hartley (1705–57) and the direct precursor of *associationism* in psychology. Hartley also proposed that sensations were paralleled by vibrations . . . or “elemental” particles in the nerves and brain providing the basis for physiological psychology. Thomas Reid (1710–96) developed Locke's idea of *inner sense* to postulate that the mind contained a number of innate faculties. It was from these faculties that Franz Joseph Gall (1758–1828) derived his list of “powers of the mind” that he attempted to localize in the brain.

However, while the British empiricists were laying the foundation for a science of psychology, Immanuel Kant (1724–1804) was denying that such a science was possible. Kant pointed out that the scientific method requires the use of mathematics and experimentation. He considered that mathematics could not be applied to the description of mental phenomena because these phenomena vary in only one dimension – time. Likewise, experimentation could not be applied to psychology because mental phenomena are private and therefore inaccessible to experimental manipulation. If we accept Kant's ideas, then physiology (the study of the brain) is a scientific discipline, while psychology (the study of the mind) is not. As a result of this distinction psychology was long considered not to be a proper subject for scientific enquiry, especially when restricted to the study of subjective experience. Even today, many traces of this unfortunate notion remain. For example, one of the many websites we consulted in the course of writing this chapter names people who have had an important role in the study of consciousness. The names are presented in three lists headed: Philosophers, Psychologists, and Scientists. Furthermore, a very eminent academic colleague of the authors recently informed us that he welcomed the advent of brain imaging since this technique would permit an objective (i.e., physiological) measure of happiness.

The Scientific Study of the Mental in the Nineteenth Century

The development of the methods of psychophysics in the nineteenth century can be seen as a reaction against the idea that mental phenomena are not amenable to experimental study and mathematical modeling. The key figure in the development of psychophysics was Gustav Fechner (1801–87). Fechner believed, against Descartes, that mind and body were two aspects of a single entity. He also believed, against Kant, that mental processes could be measured. His method of psychophysics (Fechner 1860) built on the demonstration by Herbart (1824) that mental experiences (sensations) vary in intensity and that there is a threshold (or *limen*) such that below a certain stimulus intensity there is *no* sensation. Fechner also built upon Weber's concept of the *just noticeable difference* (JND) (Weber 1834). The JND is the smallest increase in stimulus intensity that is required to produce a *change* in sensation. Fechner used the JND as the unit of measurement and showed that there was a systematic relationship between JNDs (a subjective measure of sensation) and intensity of the physical signal. Across many modalities he found that the relationship between physical stimulus intensity and subjective sensation was logarithmic (the Weber-Fechner law). He speculated that the relationship between intensity of sensation and nervous activity would also be logarithmic, but had no way of measuring nervous activity. Fechner succeeded in showing that the mental could be measured and was closely linked to the physical. He also developed some of the basic methods of experimental psychology that we still use today.

Helmholtz's Unconscious Inferences

In parallel with the emergence of experimental psychology great advances were made in the understanding of the nervous system. A key figure in this development was Hermann Helmholtz (1821–94, enobled to von Helmholtz in 1882). Helmholtz began his studies of physiology with Johannes Müller. Like most biologists of his day, Müller was a vitalist who believed that living processes could never be reduced to the mechanical laws of physics and chemistry. Life depended on a vital force that was not susceptible to experimental investigation. In particular, he believed that the nerve impulse was a vital function that could never be measured experimentally since it was not extended in time. With proper disdain for the beliefs of his PhD supervisor, Helmholtz developed the myograph and measured the speed of travel of nerve impulses. He found that this was rather slow (~27 meters per second). The slow speed of travel of nerve impulses raised the possibility that mental processes might also be slow enough to measure, a possibility that led Donders to develop the reaction time task (see later).

Helmholtz made a particular study of the neural basis of perception (Helmholtz 1866). Müller had made the important observation (which he called *the law of specific nerve energies*) that sense organs cause the same subjective experience however they are stimulated. A mechanical blow to my eye, a stimulation that has nothing to do with light, nevertheless causes me to "see stars." Müller proposed that there were specific kinds of nerves associated with each sense organ that created the subjective quality associated with each modality. Helmholtz took this idea a step further and proposed that there might be different kinds of nerves supporting perception even within modalities. Since the experience of all hues can be created by mixing three primary colors, Helmholtz followed Young (1802) in proposing that there were three different kinds of nerve fiber in the human eye concerned with color. He calculated curves for the wavelength sensitivity of these three kinds of receptor. These speculations were subsequently confirmed experimentally.

Helmholtz recognized that the law of specific nervous energies implied that sensations do not provide direct access to objects, but are signs of reality that have to be interpreted. He demonstrated this clearly in relation to the perception of depth in 3-D space. There are many visual cues to the distance of objects from us. One is the disparity between the views received by the two eyes. Another is motion parallax: the observation that, when we are moving, nearby objects move across our eye much faster than objects that are far away. Helmholtz realized that, in order to create a percept from these sensory cues, the brain must make inferences based on prior knowledge. He concluded that perception depends upon *unconscious inferences*; unconscious because our experience of perception is that it is immediate. We are not aware of the inferences being made. Through his concept of unconscious inferences Helmholtz was anticipating the idea of the cognitive unconscious that became a key feature of cognitive psychology 100 years later. He was also anticipating the recent idea of perception as Bayesian inference (Kersten, Mamassian, & Yuille 2004). The idea that inferences can be made unconsciously was controversial and Helmholtz subsequently regretted using this term. "Recently I have refrained from using the phrase *unconscious inference in* order to avoid confusion with what seems to me a completely obscure and unjustified idea which Schopenhauer and his followers have designated by the same name." (Helmholtz 1878). He presumably had in mind Schopenhauer's claim that the will is largely unconscious and manifests itself in sexual desire. But there were additional reasons for the controversy. Making inferences is an example of the rational decision-making that Descartes proposed was the preserve of the soul. By taking decisions away from the soul and assigning them to the brain, Helmholtz seemed to be undermining the idea of personal responsibility, which many people continue to believe is the basis of moral behavior. Similar arguments continue today in relation to free will and the brain (e.g., Wegner 2002; Schurger, chapter 49).

Early Progress in Physiology and Psychology

By the end of the nineteenth century much had been learned about the brain. Nerve fibers had been identified as extensions of nerve cells. This paved the way for Ramon y Cajal to propose the neuron doctrine, the idea that the nerve cell is the basic unit of the nervous system (Jones 1994). Helmholtz's fellow student, du Bois-Reymond, had demonstrated the electrical basis of nerve impulses, leading to the idea that it was energy rather than motion that was transmitted through neurons (Du Bois-Reymond 1848). Ferrier and others had located motor and sensory regions in the brain and Korbinian Brodmann had begun to identify the discrete brain regions that still bear his name (Brodmann 1909).

At the same time psychology had been established as a scientific discipline and in 1879 Wilhelm Wundt had founded the first psychology laboratory in Leipzig. Reaction time had been established by Frans Donders (1818–89) as an important technique for measuring the duration of mental events. Donders found that simple reaction times (one stimulus to which one possible response is required) were always faster than choice reaction times (two stimuli and two possible responses). He proposed that this difference reflected the purely mental process of making a choice (Donders 1868). This "subtraction" method for isolating correlates of mental processes later became the standard procedure in functional brain imaging (Posner et al. 1988).

Wundt and other early psychologists used the reaction time method extensively, but very differently from the way it is used today. Their emphasis was very much on the first-person perspective. They wanted to measure pure apperception time (the time it takes to perceive something for what it is) by subtracting away the motor response time. Participants were instructed to move in response to a stimulus and their reaction times were measured. In one condition participants were instructed to attend to the movement to be executed. This condition gave a measure of the motor response time (or muscular reaction time). In the other condition participants were asked to attend to the sense impression received from the stimulus (sensorial reaction time). The sensorial reaction time was supposed to be longer than muscular reaction time because the apperception time was added onto the motor time. In practice, the results were very variable and many subjects simply could not do the task (Cattell 1893). Great introspective skill is required to decide when a stimulus has been fully perceived.

The dominant figure in psychology at the end of the nineteenth century was William James (1842–1910), whose two-volume textbook, *Principles of Psychology*, is still well worth reading today. James identified consciousness with the stream of thought. He recognized the power of attention to give a focus and a margin to consciousness. He also recognized the importance of unconscious processes.

Psycho-Physical Processes in Attention

Given all these advances, everything was in place for renewed attempts to speculate about the neural correlates of consciousness. One such speculation comes from an article in *Brain* (1890), in which James Sully of University College London considers "[p]sycho-physical processes in attention." Three commentaries on this article appeared in a later issue of the journal. The paper is about the neural correlates of selective attention. The discussion makes an interesting comparison with discussions on the same topic over 100 years later.

In most cases of selective visual attention there is an obvious motor factor in that we move our eyes to fixate the attended object. However, Sully recognized the importance of covert attention. Once again it was Helmholtz who had pointed out this phenomenon. "It is a curious fact . . . that the observer may be gazing steadily . . . yet at the same time he can concentrate his attention on any part of the dark field he likes." In the case of covert attention, Sully asks "where is the motor factor?" In his commentary, Alfred Fouillée concludes that the answer "lies in the liberation of cerebral energy upon the sensory centers of vision, not upon the ocular muscles. Certain parts of cerebral cortex are excited, others are inhibited." Today the same ideas would be expressed with phrases such as “top-down modulation of early visual areas” and “biased competition.” Attempts to discuss the neural correlates of selective attention in 1890 suffered from two major disadvantages. First, nervous activity could be described only in terms of energy. The idea that neurons could transmit and store *information* was yet to be developed (see later). Second, experimental studies of attention emphasized subjective experience rather than behavior. Researchers were concerned to explore the experience of the act of attending and its consequences. “We are conscious of the *starting* of the centrifugal (i.e., top-down) current at the instant it is liberated by the brain” (the effort of will). “The effect of this current is to make the attended object appear more vivid” (Sully 1890).

The behaviorist school arose in part because of the difficulty and unreliability of this experimental study of subjective experience. Through their emphasis on the study of animals, the behaviorists identified markers of mental processes that did not depend upon verbal reports. The unintended legacy of behaviorism is that we now have many experimental techniques that provide robust, objective markers of conscious and unconscious processes.

Developments in the Early Twentieth Century

This period is sometimes represented as a desert as far as consciousness studies are concerned, but this is an exaggeration. It is true that John B. Watson tried to eliminate both reference to consciousness and use of introspective methods from psychology, but he did not succeed, even in the United States. Woodworth’s introductory textbook of psychology, which remained in print from 1921 to 1947, was subtitled *a study of mental life.* Stanley S. Stevens, while avoiding mentalistic language, continued the psychophysical program of research started by Fechner (Stevens 1936). Of course, psychophysics depends fundamentally upon introspection. Edward C. Tolman criticized the idea that behavior could be fully explained by chains of stimulus-response associations and proposed that both humans and rats used internal perceptual representations (cognitive maps) to guide their behavior (Tolman 1948).

In Europe, Piaget studied the development of mental processes. Bartlett studied mental processes in long-term memory. The Gestalt psychologists studied the mental processes that underlie perception. The slogan of the Gestalt psychologists, “The whole is more than the sum of its parts,” implied that complex dynamic interactions in the nervous system were fundamental to conscious experience. Of particular interest for later studies of the neural correlates of consciousness are the various perceptual illusions in which subjective experience is decoupled from physical stimulation. Many such illusions, including binocular rivalry, had already been described in the nineteenth century. However, the Gestalt psychologists emphasized the importance of these phenomena for understanding the mechanisms of perception.

The key development in the early twentieth century was the introduction of information theory by Hartley (1928) and Shannon and Weaver (1949). This is a mathematical technique that allows the amount of information in a signal, the rate of transmission of information through a communication channel, and the capacity of a communication channel to be quantified. The development of information theory was the first step in a mathematical account of cognition. If we consider information to lie in the realm of the mental rather than the physical, then information theory is also the first step in solving the difficult problem of bridging the mental and the physical domains. It is important to note, however, that the information in a signal is not the same as the meaning of a signal. Computers can transmit information but whether that information is meaningful depends on whether the receiver can interpret it.

It was immediately recognized that the brain could be treated as a communications system that processes and transmits information, rather than motion or energy. Conceiving of the brain in this way allowed the realization that it was now possible to develop intelligent machines. McCulloch and Pitts (1943) updated the neuron doctrine to state that the neuron was not simply the basic anatomical unit of the central nervous system (as Cajal had proposed) but the basic *information processing* unit. McCulloch and Pitts also proposed that the brain could be modeled by artificial neural nets constructed from very simple information processing units.

The Last 50 Years: The Triumph of Cognitive Psychology

Information theory had an immediate impact on psychology. Hick (1952) applied information theory to choice reaction time and showed that response time was directly proportional to the amount of information in the signal (i.e., log of the number of choices). Miller (1956) applied information theory to psychophysical judgments and showed that there was an upper limit (~2.6 bits, i.e., seven, plus or minus two items) to the number of categories that could be handled. He also showed that there was an upper limit for the capacity of immediate memory, but that this limit was determined by the number of items (or chunks), not by information. This approach rapidly led to the development of cognitive psychology in which psychological processes are described in engineering terms (Kenneth Craik's *The Nature of Explanation* also had a key role in this development) taken from communication theory (e.g., channel capacity), control systems theory (e.g., feedback), and computing (e.g., central processor, response buffer) (e.g., Broadbent 1958). Psychologists began to use "box and arrow" diagrams, flow charts of systems in terms of processes and information transmission.

While cognitive psychologists tended not to use the word "consciousness," this was nevertheless frequently the object of their study. Following James, the contents of "working" or "active" memory as studied by Alan Baddeley and colleagues (Baddeley 1986), can be equated (roughly) with the contents of consciousness. Deploying selective attention, as in Broadbent's dichotic listening task and Posner's covert spatial attention task (Posner 1978), requires a voluntary effort. However, cognitive psychologists tended not to use introspection as a direct source of data. Intuitions derived from introspection had to be confirmed by behavioral data. For example, introspection suggests that, after reading a telephone number, we maintain our consciousness of that number in working memory by saying it to ourselves. This implies that the visual material has been converted to an auditory representation. This intuition was confirmed when Conrad showed that confusion errors were better predicted by auditory rather than visual similarity even though the numbers had been presented visually (Conrad 1962).

The Cognitive Unconscious

Perhaps the major development for consciousness research during the past 50 years has been the demonstration of unconscious, automatic psychological processes in perception, memory, and action, named the *cognitive unconscious* by John Kihlstrom (1987).

The term subliminal perception, for example, describes the situation where the presentation of a stimulus affects subsequent behavior of the observer even though the stimulus never enters the consciousness of the observer (see Kouider and Faivre, chapter 39). In the 1960s, claims about subliminal perception were dismissed by experimental psychologists on the basis of methodological inadequacy, but the development of more sophisticated experimental techniques, such as priming (Marcel 1983) and analytic techniques such as signal detection theory (Swets, Tanner, & Birdsall 1961) provided convincing evidence. Such unconscious psychological processes were observed in more exaggerated form in patients with brain damage. Some patients with lesions in visual cortex can make correct "guesses" about the properties of visual stimuli that they cannot "see" (Weiskrantz & Warrington 1975; see Kouider and Faivre, chapter 39). Patients with dense amnesia can retain knowledge about stimuli they have no memory of having seen before (Warrington & Weiskrantz 1968). Patient DF, with damage to inferior temporal cortex, can use visual information of which she is unaware to guide her movements (Goodale et al. 1991; see Goodale, chapter 46. More recently, social psychologists have demonstrated that a whole range of unconscious processes influence social behavior (Bargh & Chartrand 1999).

The problem for psychological studies of unconscious processes is that we need a marker that such processing has taken place, but at the same time we do not want to draw the subject's attention to the stimulus that they are unconsciously processing (Mack & Rock 1998). The subject can tell us that they did not see a stimulus, but to know that they have nevertheless processed it we need additional markers, for example facilitation or interference with the processing of subsequent stimuli of which they are aware. The development of brain imaging techniques has provided additional markers of such unconscious processing. Using these techniques, we can ask if unconscious processing is associated with a specific pattern of brain activity. For example, Beck et al. (2001) showed that undetected faces in a change blindness paradigm elicited activity in fusiform cortex (see Rees and Frith, chapter 42).

Many now believe that most of the processing undertaken by the brain occurs without our awareness (Velmans 1991), but many have found the term "cognitive unconscious" confusing. This confusion results from a shift in the meaning of the word "cognitive." Previously the term cognitive (as in the term cognitive therapy) referred to knowledge, beliefs, and attitudes, all key components of consciousness. Furthermore, following Kant, sharp distinctions were made between cognition (to do with knowledge), emotion (to do with feelings), and conation (to do with will). Today, following Neisser's 1967 book *Cognitive Psychology*, many use *cognitive* (as in the terms cognitive psychology and cognitive neuroscience) to replace the older term "information processing" and to refer to what the brain does. An account of a psychological or a neural system that included a box and arrow diagram involving representations, transformations, and information flow would be called a cognitive account. From this point of view cognitive processes exist in the computational domain that lies between neural activity on the one hand and behavior and conscious experience on the other hand. Such cognitive processes need not lead to consciousness and can be evoked to explain feeling and will as well as knowledge.

The demonstration of unconscious processes raises a new problem for the study of consciousness. Just because subjects can detect or discriminate a stimulus, does not mean that they are conscious of it. Their success may be the result of unconscious processes. From their first person perspective they are just guessing.

Introspection, Protocol Analysis, and Meta-cognition

While introspection was the method of choice for nineteenth-century psychologists, this method was used far less in the twentieth century. It was not abandoned completely, however. In particular it was used in the study of problem solving. In order to gain access to the conscious processes used to solve a problem subjects were asked to "think aloud." Indeed, the arch-behaviorist John B. Watson was a pioneer in the use of this method. "The present writer has often felt that a good deal more can be learned about the psychology of thinking by making subjects think aloud about definite problems, than by trusting to the unscientific method of introspection" (Watson 1920). For Watson thinking aloud was not introspection, but verbal behavior. However, it is not clear to us what someone "thinking aloud" is doing, if not introspecting. The method was used extensively by Duncker (1945), one of the Gestalt psychologists, and refined as "protocol analysis" by Ericsson and Simon (1984). Nevertheless, methodologies for harnessing introspection as a source of data have lagged behind those developed for behavioral tasks. In recent years there has been increasing interest in developing such methods (Jack & Roepstorff 2004).

Thinking aloud is a form of meta-cognition since subjects must reflect upon and report their thoughts. Meta-cognition has been used in a clever way to provide behavioral measures that reflect consciousness and hence a first-person perspective. For example, to make the confidence ratings used in psychophysics experiments, subjects must think about their perceptions. If the degree of confidence correlates with the accuracy of the judgments then we can conclude that the subjects were conscious of the stimuli rather than just guessing (Kunimoto, Miller, & Pashler 2001). This approach has been used in the study of animal consciousness. Monkeys can be trained to make confidence judgments and these behavioral responses can be used as evidence of whether or not they are conscious of stimuli (Cowey & Stoerig 1997; Hampton 2001).

The same idea underlies the process dissociation technique developed by Jacoby (1992). Subjects are asked to decide whether a word was previously presented in list A rather than list B. The assumption is that subjects can reject a familiar word from list A only if they can consciously recollect that it was in list B. Here again a behavioral response is being driven by introspection.

The Current State of Consciousness Research

Despite much progress consciousness remains as elusive as ever. Some difficulties have been resolved, but new ones have emerged. At the beginning of the nineteenth century, there was little distinction between consciousness and life itself, with both depending upon vital essences that were not amenable to experimental study. The monster created by Frankenstein in Mary Shelley's novel has not only life, but also an exquisite sensitivity to human experience and suffering. Science gradually dispelled the need for vital essences to explain life, but consciousness remained unexplained. By the early twentieth century, in James Whale's version, the monster lives, but is only dimly conscious. By the end of the century the monster has evolved into a plague of zombies who behave like humans (Horne 1992), while having no consciousness.

Zombies retain a surprisingly strong influence on contemporary philosophers of consciousness. They (that is the philosophers) are interested in the existence of a particular kind of zombie, which is physically and behaviorally identical to us, but is not conscious. Neuroscientists and psychologists, in contrast, are interested in a form of Haitian zombie that is not conscious, but in which the cognitive unconscious (the zombie within) is intact (Koch & Crick 2001). In what way would such a creature be distinguishable from us?

In the twenty-first century, we know that life does not depend upon a vital essence, but we are still not sure about consciousness. Perhaps there is a vital essence that turns a zombie into a human. There are various proposals as to the nature of this vital essence. Eliminative materialists (e.g., Paul and Patricia Churchland) have concluded that consciousness is itself a vital essence and therefore does not really exist (see Mandik, chapter 33). For *functionalists*, following in the footsteps of La Mettrie, the vital essence is a computational algorithm of sufficient complexity. This can be instantiated in silicon just as well as in neurons. If a machine has the right kind of complexity it will be conscious. No new physical principles will be required to understand how it works (see Aleksander, chapter 7). Others claim that some as yet undiscovered scientific process, such as quantum entanglement at a macroscopic level, is needed to explain consciousness (e.g., Stuart Hameroff, see Atmanspacher, chapter 21). And finally *mysterians* think that the problem of consciousness is so complex that the human brain can never explain it (e.g., Colin McGinn, see McLaughlin, chapter 30).

Meanwhile the scientific study of mental processes has revealed that consciousness is not necessary for rational thought. Inferences can be drawn and decisions made without awareness. This raises a new problem for our understanding of consciousness. Descartes and his contemporaries took it for granted that consciousness was necessary for rational thought and willed, as opposed to automatic, behavior. If not the basis of rational thought, what is the function of consciousness? Again extreme positions have been taken up. On the one hand, consciousness is considered to have no function. It is just an epiphenomenon, which can have no impact on the physical world (see Kim, chapter 32; and Schurger, chapter 49). On the other hand, the followers of Darwin claim that consciousness has evolved and must therefore give some advantage to those of us who have it (see Polger, chapter 6). From this perspective the sophisticated forms of consciousness found in humans may be associated with language and the creation of culture. Perhaps consciousness is necessary for communicating mental states and sharing experiences? This is not a new idea. Nietzsche made the conjecture "that consciousness in general developed itself only under the pressure of the need to communicate."

Consciousness studies are frequently criticized for failing to define precisely what consciousness is. In this respect there has been little change over the past two centuries. In part the problem arises because consciousness remains a common-sense term rather than a scientific one. Different people use the term to mean different things (see Tye, chapter 2). Studies purporting to define the neural correlates of consciousness often address only one aspect of consciousness (e.g., access consciousness) while leaving other aspects (e.g., phenomenal consciousness) untouched. A likely consequence of the intellectual endeavors promoted in this book is that this fractionation of consciousness will become more explicit and the different components associated with specific operational definitions. In the final section of this introduction we describe some specific problems in the study of consciousness, which, when answered, will aid the development of such fractionations and definitions.

Scientific Questions

The historical developments that we have charted in this chapter have profoundly shaped current thinking about the outstanding major scientific questions concerning consciousness. Many of these questions, particularly those concerning the cognitive and neural basis of consciousness, could not have been asked even 20 years ago. These are not questions about the really hard problems of consciousness (see Chalmers, chapter 3). Rather they are questions for which satisfactory answers will soon be found. When they are answered the hard problems may seem easier.

A. Are there different kinds of consciousness?
A major section of this book is devoted to varieties of consciousness, so the answer to this question must be affirmative (see the section on *Some Varieties of Conscious Experience*; also Tye, chapter 2). However, we neither know the precise fractionation of consciousness, nor yet know in what way these different kinds of consciousness will vary. Are the differences simply quantitative, with dreaming, fringe consciousness and core consciousness being just simplified versions of waking, focal, and self-consciousness? Or are there qualitative differences between these different kinds of consciousness? These questions about the varieties of consciousness can be answered through studying the cognitive and neural correlates of the different varieties of consciousness (as well as introspective reports). Are certain representations and computations only possible for certain kinds of consciousness? Are different patterns of neural activity associated with different kinds of consciousness? The questions can also be addressed by contrasting the consciousness of animals and humans (Allen & Trestman, chapter 5), or the consciousness of infants and adults (Trevarthen & Reddy, chapter 4).

B. Are there biological markers of consciousness?
This question has been dramatically sharpened by the demonstration of multiple unconscious processes. We can now ask about the differences between those processes that are associated with consciousness and those that are not (see the section on *Cognitive Psychology of Consciousness* and chapters by Baars (16), Kouider & Faivre (39), Kihlstrom, Dorfman, & Park (40), and Rees & Frith (42)). Do the processes associated with consciousness involve specific kinds of computations and representations? Are they associated with specific kinds of neural activity, and do they involve particular regions of the brain? By contrasting conscious and unconscious processes we already know, for example, that activity in a region of human fusiform cortex is necessary, but not sufficient for the conscious experience of a face.

C. How do we determine the presence of consciousness?
This remains an intensely practical question that confronts clinicians in the intensive therapy unit and the operating theater (see Kihlstrom & Cork, chapter 48). Is this brain-damaged patient in a coma (i.e., unconscious) or are they instead in a locked-in state: conscious of everything that is being said, but unable to move any part of their body? Evidence of consciousness is currently inferred behaviorally, but does the resulting classification of patients into coma, minimally conscious, persistent vegetative state, or locked-in syndrome accurately reflect the underlying degree of consciousness of such patients?

Precisely the same problem confronts anesthetists daily. The patient on the operating table cannot move because they have been injected with a muscle relaxant, but if they become conscious in the middle of the operation they will experience distress and perhaps even sue the hospital. How can the anesthetist tell if their patient is awake? The solution is to find reliable neural correlates of consciousness, or to find some way of communicating with the patient. But how do we determine consciousness when high-level communication is not available, as with animals, infants, or machines? Neural markers of consciousness may be relevant for determining consciousness in animals and infants with brains, but is not relevant for most machines. Is there some cognitive process that is a marker of consciousness?

D. What is consciousness for?
The demonstration of unconscious processes has also sharpened our thinking on this question. We can ask whether there is some kind of problem that can be solved by conscious processes, but not by unconscious ones. In other words, although Hollywood zombies can go shopping (Romero 1978), are there other tasks that they find more difficult, or cannot perform? Various candidates have been proposed, for example, the analysis of complex or novel input, the operation of working memory, learning of novel material, thinking and planning, speech production and reading, and the performance of any task that is novel, or that requires flexibility and feedback.

The reader will have noticed that all these questions are closely inter-related. Determining if someone is conscious will depend upon finding markers of consciousness. Finding cognitive markers of consciousness may give clues about what consciousness is for. Alternatively, if we knew what consciousness was for, then it might be easier to find markers of consciousness, and so on. More importantly, by the end of this book, the reader should be convinced that these are questions we are now in a better position to answer.

See also 2 Philosophical problems of consciousness; 3 The hard problem of consciousness; 42 Methodologies for identifying the neural correlates of consciousness.

Further Readings

Hilgard, E. R. (1980) Consciousness in contemporary psychology. *Annual Review of Psychology* 31, 1–26.
Velmans, M. (2009) *Understanding Consciousness*. London: Routledge/Psychology Press, chs. 1 to 5.

References

Baddeley, A. (1986) *Working Memory*. Oxford: Oxford University Press.
Bargh, J. A. and Chartrand, T. L. (1999) The unbearable automaticity of being. *American Psychologist* 54: 7, 462–79.
Beck, D. M., Rees, G., Frith, C. D., and Lavie, N. (2001) Neural correlates of change detection and change blindness. *Nature Neuroscience* 4: 6, 645–50.
Broadbent, D. E. (1958) *Perception and Communication*. London and New York: Pergamon Press.

Brodmann, K. (1909) *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Principien, dargestellt auf grund des Zellenbaues*. Leipzig: Johann Ambrosius Barth.
Cattell, J. M. (1893) Aufmerksamkeit und Reaction. *Philosophische Studien* 8, 403–6.
Conrad, R. (1962) An association between memory errors and errors due to acoustic masking of speech. *Nature* 193, 1314–15.
Cowey, A. and Stoerig, P. (1997) Visual detection in monkeys with blindsight. *Neuropsychologia* 35: 7, 929–39.
Donders, F. C. (1868) On the speed of mental processes. Translated by W. G. Koster 1969. *Acta Psychologica* 30, 412–31.
Du Bois-Reymond, E. (1848) *Untersuchungen über thierische Elektricität*. Berlin: Georg Reimer.
Duncker, K. (1945) *On Problem Solving*. Washington, DC: American Psychological Association.
Ericsson, K. A. and Simon, H. (1984) *Protocol Analysis: Verbal Reports as Data*, Cambridge, MA: MIT Press.
Fechner, G. T. (1860) *Elemente der Psychophysik*. Leipzig: Breitkopf und Härtel.
Goodale, M. A., Milner, A. D., Jakobson, L. S., and Carey, D. P. (1991) A neurological dissociation between perceiving objects and grasping them. *Nature* 349: 6305, 154–6.
Hampton, R. R. (2001) Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences of the United States of America* 98: 9, 5359–62.
Hartley, R. V. L. (1928) Transmission of information. *Bell Systems Technical Journal* 7, 535.
Helmholtz, H. v. (1866) *Handbuch der Physiologischen Optik*. Leipzig: Voss.
Helmholtz, H. v. (1878) The facts of perception. *Selected Writings of Hermann Helmholtz*. Middletown, CT: Wesleyan University Press.
Herbart, J. F. (1824) *Psychologie als Wissenschaft, neu gegründet auf Erfahrung, Metaphysik und Mathematik*. Königsberg: A. W. Unzer.
Hick, W. E. (1952) On the rate of gain of information. *Quarterly Journal of Experimental Psychology* 4, 11–26.
Horne, P. (1992) I shopped with a zombie: consumer culture, fiction and cinema. *Critical Quarterly* 34: 4, 97–110.
Jack, A. I. and Roepstorff, A. (2004) *Trusting the Subject?* Exeter, UK: Imprint Academic.
Jacoby, L. L. (1992) A process dissociation framework: separating automatic from intentional uses of memory. *Journal of Memory and Language* 30, 513–41.
Jones, E. G. (1994) The neuron doctrine. *Journal of History of Neuroscience* 3, 3–20.
Kersten, D., Mamassian, P., and Yuille, A. (2004) Object perception as Bayesian inference. *Annual Review of Psychology* 55, 271–304.
Kihlstrom, J. F. (1987) The cognitive unconscious. *Science* 237: 4821, 1445–52.
Koch, C. and Crick, F. (2001) The zombie within. *Nature* 411: 6840, 893.
Kunimoto, C., Miller, J., and Pashler, H. (2001) Confidence and accuracy of near-threshold discrimination responses. *Consciousness and Cognition* 10: 3, 294–340.
Mack, A. and Rock, I. (1998) *Inattentional Blindness*. Cambridge, MA: MIT Press.
Marcel, A. J. (1983) Conscious and unconscious perception: experiments on visual masking and word recognition. *Cognitive Psychology* 15: 2, 197–237.
McCulloch, W. S. and Pitts, W. H. (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5, 115–33.
Miller, G. A. (1956) The magical number seven, plus or minus two. *The Psychological Review* 63, 81–97.
Posner, M. I. (1978) *Chronometric Explorations of Mind*. Oxford: Lawrence Erlbaum.

Posner, M. I., Petersen, S. E., Fox, P. T., and Raichle, M. E. (1988) Localization of cognitive operations in the human brain. *Science* 240: 4859, 1627–31.
Romero, G. A. (1978) *Dawn of the Dead* (Italy/USA, United Film Distribution Company).
Shannon, C. E. and Weaver, W. (1949) *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
Stevens, S. S. (1936) A scale for the measurement of a psychological magnitude: loudness. *Psychological Review* 43, 405–16.
Sully, J. (1890) The psychophysical process in attention. *Brain* 13: 2, 145–64.
Swets, J. A., Tanner, W. P. J., and Birdsall, T. G. (1961) Decision processes in perception. *Psychological Review* 68, 301–40.
Tolman, E. C. (1948) Cognitive maps in rats and men. *Psychological Review* 55: 4, 189–208.
Velmans, M. (1991) Is human information processing conscious? *Behavioural and Brain Sciences* 14: 4, 651–68.
Warrington, E. K. and Weiskrantz, L. (1968) New method of testing long-term retention with special reference to amnesic patients. *Nature* 217: 132, 972–4.
Watson, J. B. (1920) Is thinking merely the action of language mechanisms? *British Journal of Psychology* 11: 8, 7–104.
Weber, E. H. (1834) *De pulsu, resorptione, auditu et tactu. Annotationes anatomicae et physiologicae*. Leipzig: Koehler.
Wegner, D. (2002) *The Illusion of Conscious Will*. Cambridge, MA: Bradford Books, MIT Press.
Weiskrantz, L. and Warrington, E. K. (1975) Blindsight – residual vision following occipital lesions in man and monkey. *Brain Research* 85: 1, 184–5.
Young, T. (1802) On the theory of light and colours. *Philosophical Transactions of the Royal Society of London* 92, 12–48.