

# Consciousness Studies

Joseph A. Goguen

Department of Computer Science and Engineering  
University of California at San Diego, La Jolla CA 92093-0114

Consciousness studies is a new, rapidly evolving, highly interdisciplinary field. Disciplines involved include psychology, philosophy, physics, sociology, religion, dynamic systems, mathematics, computer science, neuroscience, art, biology, cognitive science, anthropology, and linguistics. Even in the early 1990s, most scientists considered consciousness taboo, but now many consider it the most important unsolved problem in science.

Consciousness is also a key issue in the ongoing dialogue between science and religion. The dominant view of consciousness in the hard sciences is of course materialist and reductionist. This has had important successes, but it also has important unresolved problems. For example, the biologist Francis Crick wrote, “You’re nothing but a pack of neurons” [Crick, 1994] in parody of Lewis Carroll. But most people, including those in consciousness studies, and even most in neuroscience, feel there is much more to human life than can be seen at the level of neurons.

Notions of consciousness are important in many religions. The term “God consciousness” is important in the protestant theology of Friedrich Schleiermacher [Schleiermacher, 1989] and his followers, and “Christ consciousness” is used in some Christian and New Age religions, sometimes in a doubtful way. “Cosmic consciousness” is important in Hinduism, especially Vedanta, and “pure consciousness” is important in the Buddhist school called *Dzogchen* in Tibetan, and *Maha Ati* or *Mahasandhi* in Sanskrit; consciousness is a common theme in the Tantric traditions. Reports of meditation experience are taken more seriously in consciousness studies than in the hard sciences, which often dismiss such data as “mere subjective experience.” On the other hand, due to close connections with various religions, some writers on consciousness have hidden (or not so hidden) agendas, so that caution is called for.

In general, we may say that the hard sciences attempt to reduce consciousness to the material, while religions are more concerned with mental (or spiritual) aspects. This reflects the heritage of mind-body separation associated with the philosopher René Descartes. Although there is no single dominant view of consciousness, nor even any generally accepted definition, there has been substantial progress.

## Shape of the Discipline

The field has a professional society, the Association for the Scientific Study of Consciousness (abbreviated ASSC), one highly interdisciplinary journal, the *Journal of Consciousness Studies* (abbreviated *JCS*), and three journals devoted mainly to scientific and philosophical studies, *Consciousness and Cognition*, *Consciousness and Emotion*, and *Psyche*, the latter being an electronic journal; *JCS* also has a popular online discussion group. Many other journals publish papers on consciousness, such as *Behavioral and Brain Sciences* and *Mind*. There is an important biannual conference series held in Tucson, Arizona since 1994, organized by the University of Arizona, which also has an important research center on consciousness studies. *Consciousness and Cognition* and *Psyche*

are official journals of ASSC, which also organizes a biannual conference. Well known universities offering courses on consciousness include (at least) New York University, Bryn Mawr, Vanderbilt, and the Universities of Colorado, Virginia, and Arizona. Advanced degrees in consciousness studies are offered by (at least) the University of Skövde in Sweden, Greenwich University in Australia, and Birla Institute of Technology and Science in India; also John F. Kennedy University has a Department of Consciousness Studies, and Brunel University offers an MSc degree in Cognition and Consciousness. In addition, there are very many one off conferences, and the emergence of the specialized journal *Consciousness and Emotion* in 2000 seems a sign that the field is maturing.

### Issues, Paradigms, and Results

It is difficult to single out any small set of issues, not only because of the rapid growth of the field, but also because each of its many paradigms defines different sets of issues as central, secondary, marginal, and meaningless. Nonetheless, the following are some issues, paradigms, and results that seem most important in the literature.

The most obvious issue is *how* to study consciousness. Despite the fact that the advocates of various approaches are in constant dialogue, sometimes acrimonious, no approach has been completely discredited, except perhaps that of mediums, spiritualists, etc. This is why the editorial policy of *JCS* calls for a wide diversity of views, and aims to promote dialogue among them, and why the Tucson conference follows a similarly liberal policy. As the distinguished philosopher John Searle said:

At our present state of the investigation of consciousness, we *don't know* how it works, and we need to try all kinds of different ideas.

Nevertheless, journals and conferences devoted to specific aspects of consciousness studies can certainly be very valuable.

The relation between mind and body is another major issue. Are mind and body the same kind of thing, or are they different? Or perhaps the same thing but differently perceived? Monism says there is just one kind of thing, and material monism (also called physicalism) says that all things are material, while mental monism (also called idealism) says that all things are mental. The dualism associated with Descartes says that both material and mental things exist. There are many variants of these and many other positions. Reductionist scientific approaches like neuroscience and experimental psychology of course are material monist. The philosopher David Chalmers is a kind of dualist, who argues that in addition to matter, information is a second fundamental world constituent [Chalmers, 1996]. The philosopher Paul Churchland is an “eliminative materialist” monist, who argues that there is really no such thing as consciousness [Churchland, 1989]. Searle is an “emergent materialist” monist, who argues that consciousness is a distinct level of phenomena, emerging out of lower level brain activity, which “only exists when it is experienced” [Searle, 1997].

It is difficult to find adherents of either dualism or mental monism among eminent scientists. Perhaps the only prominent recent exception is the Nobel prize winning physiologist Sir John Eccles (1903–1997), who advocated a form of interaction dualism similar to that of Descartes. Bishop George Berkeley (1685–1753) was the last major Western philosopher to advocate mental monism. On the other hand, dualism is the most common position in Christianity, as is mental monism in far Eastern religions. For example, the Buddhist school of Yogacara posits a form of mental monism, and is considered foundational for Buddhist Tantra. Hinduism and Taoism can also be considered mental monist.

An amusing development in the debate among these positions concerns the possibility (or impossibility) of “philosophical zombies,” creatures having exactly the same physical structure as ordinary humans, but without consciousness (*JCS* 2, 4, 1995 contains a symposium on zombies). Metaphysical debates about basic world substances seem to contribute little to our understanding of consciousness; but reconceptualizing the two main views as the scientific and phenomenological methods instead of reifying them as world substances, leads to the more fruitful goals of refining them and of seeking productive ways to combine them.

A once dominant approach now in decline is that of early cognitive science and artificial intelligence, often called cognitivism. This paradigm has a model of the mind which identifies cognition with computation, and the brain as the hardware on which it runs. The lineage of cognitivism traces back to pioneering work of Norbert Wiener on cybernetics, and to the Macy Conferences, organized since 1947 by anthropologists Margaret Mead, Gregory Bateson, and others, introducing systems theory to a key cross-disciplinary group. But cognitivists often ignore these antecedents, and instead cite the scalding review of Burrhus F. Skinner’s 1957 book *Verbal Behavior* by the linguist Noam Chomsky [Chomsky, 1959]. Skinner advocated behaviorism, a psychological theory which tried to ignore internal mental states. Chomsky argued that such states are needed to process even simple syntax. Another seminal cognitivist work, by George Miller, Eugene Galanter and Karl Pribram [Miller et al., 1960] proposed that human plans have the same structure as a certain simple kind of computer program. This tradition generally relies on formal logical representations of knowledge about the world. The cognitivist paradigm flourished from the 1960s, partly fueled by large military funding for artificial intelligence.

Cognitivism has been much criticized. A famous early attack was Searle’s Chinese room argument [Searle, 1980], which challenged the idea that a program running on a machine could be conscious. Another serious challenge came from James Gibson’s work on affordances [Gibson, 1979], showing that many cognitive tasks are greatly simplified by relying on information already in the world, instead of complex internal representations. Work in cognitive linguistics showed that many basic metaphors rely on innate sensory-motor schemas [Lakoff and Johnson, 1980]. The sociologist Lucy Suchman [Suchman, 1987] showed that plans as actually used can have structure and execution very different from that postulated by [Miller et al., 1960]. Francisco Varela (1946–2001), Evan Thompson, and Eleanor Rosch (a biologist, a philosopher, and a psychologist, all Buddhists) used empirical evidence to argue that cognition is necessarily embodied [Varela et al., 1991], rather than disembodied like a computer; they also drew on Buddhist philosophy to show how cognition is possible without a “self.” This book also constitutes a brilliant synthesis of cognitive science and religion. Rodney Brooks of MIT built robots which demonstrate that logical representation of knowledge is not necessary for the embodied action of locomotion [Brooks, 1991]. The anthropologist Edwin Hutchins showed that real world cognition is often distributed over individuals, rather than localized in a single individual [Hutchins, 1995], one example being navigation on large ships. There is also a growing body of work showing that, rather than cognition being rational and disembodied, emotion plays a central role [Panskepp, 1998]. All these developments are deeply inconsistent with cognitivism, though the significance of work done before about 1990 was not generally appreciated at the time.

Phenomenology is an area of philosophy with important implications for consciousness; it seeks to ground everything in the actual experience of human beings, i.e., it takes a “first person” experiential perspective, rather than “third person” scientific perspective (these terms are derived from grammatical terminology for pronouns). Important names here include Edmund Husserl, Martin Heidegger, and Maurice Merleau-Ponty. Heidegger considered implications of embodiment,

including finitude and temporality, noting that we are historical beings, bounded in time, space, and ability. Many of these themes also appear in the anti-cognitivist movement. Another such theme, with origins in Heidegger and especially Merleau-Ponty, but developed by Hubert Dreyfus, is the phenomenological critique of representation, which draws on our experience with routine activities to argue that representations are not necessary for embodied action. The work of Merleau-Ponty predates Gibson and Brooks, but is non-empirical, while Dreyfus makes compelling use of work by Walter Freeman connecting brain dynamics with chaos theory [Freeman, 1991].

The decline of cognitivism has inspired a return to naturalism, the study of cognition as it actually occurs in living human beings, and in particular, a shift towards neuroscience and evolutionary biology. Neuro-reductionism is perhaps the dominant position now. Certainly we can find “neural correlates of consciousness,” i.e., patterns of neural activity that correlate with various conscious experiences, such as visual perception. But it remains unclear that such correlates can ever explain the nature of consciousness. A narrower version of this challenge is to explicate qualia, which are the qualitative aspects of consciousness, e.g., “how it feels” when one is angry, or when one sees the blue of the sky. David Chalmers has introduced an influential distinction between the “easy” and the “hard” problems of consciousness studies:

The easy problems are those of finding neural mechanisms and explaining cognitive functions: the ability to discriminate and categorize environmental stimuli, the capacity to verbally report mental states, the difference between waking and sleeping. The hard problem is that of *experience*: why does all this processing give rise to an experienced inner life at all? While progress is being made on the easy problems, the hard problem remains perplexing.

One approach to bridging this gap is to postulate that consciousness is some form of emergent activity of the brain. A familiar example of an emergent property is the liquidity of water, which arises from a sufficiently large collection of water molecules at an appropriate temperature.

Another problem with which neuroscience is struggling, is the binding problem, to determine how the brain integrates sensory input from different times and/or different modalities to create a coherent seeming whole. Few doubt that this problem is solvable within the neuro-reductionist paradigm, though the complete answer will likely be complex. Neuro-reductionism has been especially successful in studying perception, and this has inspired interesting speculations on consciousness. It is intriguing that this kind of research can be used to help understand art; for example, see the articles by Semir Zeki, Vilayanur Ramachandran & William Hirstein, and others in [Goguen, 1999]. However, critics have complained that the cultural aspects of art get short changed by neuro-reductionist analyses.

Another problem is to determine the modularity and plasticity of the brain and the mind. Studies have found brain locations associated with many mental functions, but other functions have been shown to be non-local; recent work has demonstrated physical brain change associated with learning, even relatively late in life. There is strong support for the modularity of many unconscious perceptual processes, and for the non-modularity of many higher level conscious processes. Whether there is a language module, as claimed by Chomsky, remains contentious, but there is a growing consensus against his claim that language emerged suddenly “out of the blue.”

On the interface between neurophysiology and computer science is the issue of modeling neurons, networks of neurons, and ultimately, brains. A classic 1943 paper of Warren McCulloch and Walter Pitts introduced the first such model, in which neurons were either “on” or “off,” i.e., firing or not firing. These neurons are similar to the logic gates of computers, but are far simpler than

real neurons. Some key ideas due to Donald Hebb [Hebb, 1949] include the following: connections between neurons become tighter the more they are used; neurons act in groups called cell assemblies; and cell assemblies are the basis of short term memory, but not long term memory. Although these only approximate the complex functioning of real neurons (involving numerous chemical reactions, etc.), they inspired a new generation of models having important engineering applications, such as character recognition. But because of the approximation, many researchers prefer to call their work parallel distributed processing or connectionism, rather than neural net modeling.

Meanwhile, experimental neuroscience has uncovered even more complexities, some of which may have profound implications for consciousness. Benjamin Libet showed that voluntary acts are preceded by a readiness potential (a gradual negative shift in electrical potential, as recorded at the scalp) about 550 ms before the action occurs, and about 200 ms before subjects recorded a conscious intent to act [Libet, 1985]. This has been taken to imply that consciousness is constructed well after the fact, and even that consciousness may be unimportant. Mirror neurons are another significant discovery. The Italian neurophysiologist Giacomo Rizzolatti found that certain cells in monkey frontal lobes respond to specific actions, not only in the subject, but also when the subject observes another perform that same action. It has been suggested that this can help explain many puzzles, such as how we learn by imitation, or how we can put ourselves in the place of another, in order to outsmart them. To this list, I would add compassion, the capacity to empathize with others. Blindsight is another intriguing phenomenon, in which (for example) a subject reports inability to see an object, but can still guess its location with reasonable accuracy [Weiskrantz, 1997]. This dissociation between perception and awareness raises questions about the relation between conscious and unconscious processes.

Physicists have not been shy to speculate about the relevance of quantum mechanics to consciousness. This is unsurprising, since the two have long been linked by the “Copenhagen interpretation” of Neils Bohr (as augmented by John von Neumann), which says that the consciousness of an observer is needed to “collapse” the state probability distribution associated to the wave function down to a single state, when an experiment is performed. This was always controversial, but it remains respectable. Roger Penrose, instead of explaining quantum mechanics with consciousness, seeks to explain consciousness with quantum mechanics. The results seem stimulating but disappointing, since his major conclusion is that some as yet non-existent physics (quantum gravity) is needed [Penrose, 1989]. He also argues against cognitivism, though this relies on a Platonist philosophy of mathematics, in which abstract mathematical objects are just as real as chairs, trees, and people. David Bohm is another physicist who has written about consciousness, particularly in relation to the (non-sectarian) spiritual teachings of Jiddu Krishnamurthy, which helped inspire some novel versions of quantum mechanics, having philosophical interpretations that involve information and consciousness.

Attempts to prove the objective existence of God have a long and important history. Although every such attempt has failed, the dialectic of refutation and refinement has been surprisingly productive, especially in certain areas of formal logic. This is relevant to consciousness studies because if the traditional Christian God exists, then dualism is true, the mind-body problem is solved, and the hard problem of consciousness takes on a very different, more theological, character. Using modern tools from information theory, William Dembski has attempted a sophisticated revival of an ancient proof, that God exists because an intelligent designer is needed to account for the regularities of the universe [Dembski, 1999]. Dembski’s work has been greeted with massive skepticism, and even hostility, by the scientific community, in part because he is an anti-Darwinian fundamentalist Christian, who advocates teaching creation science in the schools, and in part due

to technical difficulties in his argument.

A very different God is discussed by Anthony Freeman, motivated by the idea of treating conscious states as emergent properties of brain states [Freeman, 2001]. Freeman views both God and the soul as emergent from individuals and communities, claiming that this is neutral between dualism and reductive materialist monism. He draws on work of Schleiermacher, Searle, and recent advocates of a more social approach to biology, such as Raphael Nuñez. It seems to me that a simpler approach than that of Dembski or Freeman is to place the existence of God in the category of first person experience, rather than third person fact.

### Some Emerging Trends

Biologists are applying socio-biology and evolution to consciousness, though most results are speculative, e.g., about the possible co-evolution of language and consciousness. An interesting exception is the area of ethics; for example, see the brilliant collection in *JCS* 7, no. 1/2, 2000, edited by Leonard Katz.

There have been proposals to merge phenomenology and science (such as the neurophenomenology of Francisco Varela [Varela, 1996]), and even proposals to reformulate science based on phenomenology. We can expect to see more of these, in part because experience provides phenomena that demand explanation, including the following aspects of consciousness: it is ineffable, open, fluid, non-local, temporally thick, and involves qualia and a sense of self (can it be mere coincidence that similar properties are often attributed to God?). Reports from experienced meditators suggest additional phenomena, e.g., that advanced consciousness has no thoughts. Moreover, the emphasis on time in phenomenology resonates well with many issues and results in neuroscience.

Another approach, sometimes called “second person,” is to relate consciousness to society rather than just individuals. One example is the cultural-historical approach, in the tradition of philosophers Giambattista Vico, Wilhelm Dilthey (who built on Schleiermacher), and John Stuart Mill, and of the more recent Russian activity theory of Lev Vygotsky, Alexander Luria and others; it is also related to distributed cognition and the actor-network theory of Bruno Latour and others. The area of sociology called ethnomethodology also seems promising. The hope of second person approaches is to transcend the problematic relationship between mind and body; debates here often parallel those in consciousness studies, and emerging syntheses like the cultural psychology of Michael Cole [Cole, 1996] could likely illuminate some issues in consciousness.

Moving away from the social sciences, we should certainly expect PET and fMRI techniques to continue to yield provocative results about brain function. Also, dynamical systems and chaos theories seem like areas where more can be expected. Perhaps semiotics can also make a contribution. Areas like ecology, feminism, and literature should also be given a chance. Definitely, we can predict that there will be more fermentation, discussion, and progress.

### References

- [Brooks, 1991] Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, 47:139–159.
- [Chalmers, 1996] Chalmers, D. (1996). *The Conscious Mind: In search of a fundamental theory*. Oxford.

- [Chomsky, 1959] Chomsky, N. (1959). A review of B.F. Skinner's Verbal Behavior. *Language*, 35(1):26–58.
- [Churchland, 1989] Churchland, P. (1989). *A Neurocomputational Perspective: the nature of mind and the structure of science*. MIT.
- [Cole, 1996] Cole, M. (1996). *Cultural Psychology: A once and future discipline*. Harvard.
- [Crick, 1994] Crick, F. (1994). *The Astonishing Hypothesis: The scientific search for the soul*. Scribner & Sons.
- [Dembski, 1999] Dembski, W. (1999). *Intelligent Design: The bridge between science and theology*. Intervarsity Press.
- [Freeman, 2001] Freeman, A. (2001). God as an emergent property. *Journal of Consciousness Studies*, 8(9/10):147–159.
- [Freeman, 1991] Freeman, W. (1991). The physiology of perception. *Scientific American*, 264:78–85.
- [Gibson, 1979] Gibson, J. (1979). *An Ecological Approach to Visual Perception*. Houghton Mifflin.
- [Goguen, 1999] Goguen, J., editor (1999). *Art and the Brain*. Imprint Academic. Special issue of *Journal of Consciousness Studies*, volume 6, No. 6/7.
- [Hebb, 1949] Hebb, D. (1949). *The Organization of Behavior: A neuropsychological theory*. Wiley.
- [Hutchins, 1995] Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.
- [Lakoff and Johnson, 1980] Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By*. Chicago.
- [Libet, 1985] Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8:529–566.
- [Miller et al., 1960] Miller, G., Galanter, E., and Pribram, K. (1960). *Plans and the Structure of Behavior*. Holt, Rhinehart & Winston.
- [Panskepp, 1998] Panskepp, J. (1998). *Affective Neuroscience*. Oxford.
- [Penrose, 1989] Penrose, R. (1989). *The Emperor's New Mind*. Oxford. Vintage paperback edition, 1990.
- [Schleiermacher, 1989] Schleiermacher, F. (1822/1989). *The Christian Faith*. T & T Clark.
- [Searle, 1980] Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3:417–457.
- [Searle, 1997] Searle, J. (1997). *The Mystery of Consciousness*. Granta Books.
- [Suchman, 1987] Suchman, L. (1987). *Plans and Situated Actions: The Problem of Human-machine Communication*. Cambridge.
- [Varela, 1996] Varela, F. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, 3(4):330–349.

[Varela et al., 1991] Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind*. MIT.

[Weiskrantz, 1997] Weiskrantz, L. (1997). *Consciousness Lost and Found*. Oxford.