From Cognitive Science to A Science of Consciousness

B. Alan Wallace

In this essay I intend to demonstrate how cognitive science, which stands at the crossroads of the natural sciences and the human sciences, has adopted an "objectivist" perspective on cognition that unnecessarily limits our understanding of the human mind; and I shall conclude with a prolegomenon to understanding the nature of consciousness and its causal efficacy. The essay consists of four parts: (1) comments on the historical origins of the objectivist perspective, (2) an analysis of cognitivism, (3) an analysis of connectionism, based largely on the book *The Embodied Mind: Cognitive science and human experience* by neuroscientist Francisco Varela, philosopher Evan Thompson, and cognitive psychologist Eleanor Rosch (MIT Press, 1991), and (4) suggestions for freshly examining the nature and causal efficacy of consciousness.

On the Origins of Objectivism in Modern Science

The scientific revolution of the 16th-17th centuries brought with it a new mechanical and experimental philosophy that broke away from Aristotelian doctrines and from certain dogmas of Christian theology. This new philosophy also formed its characteristic doctrines in contradistinction to the tenets of the various forms of natural magic that were promulgated during this same era. The cosmos of the natural magicians was one that was pervaded by a world soul, populated and influenced by other spiritual entities, such as angels, spirits and demons, and saturated with occult properties whose discovery awaited the investigations of natural magic. This was a participatory universe, in which the human mind, and especially the imagination, could influence not only one's own body but other physical objects and events. Human consciousness and other spiritual agencies thus participated in *enacting* the universe, in which humanity played a significant role.

Descartes formulated a mechanical philosophy that stood in direct opposition to natural magic. In the universe as he conceived it, virtually all natural processes could be understood purely in terms of inert matter, utterly devoid of influence by any spiritual agency. Although God was regarded as the Creator of the universe, He no longer

influenced the course of natural events; all influence by other spiritual entities, such as angels and demons, was denied; and even animals were regarded as unconscious automata, devoid of subjective states of feelings, desires and awareness. In all of nature, only humans retained consciousness. But the human mind was not a natural phenomenon; rather it was an immortal soul that came as a direct gift from God and supernaturally influenced the body via the pineal gland. In all of nature, then, only the human body was causally influenced by awareness. Leibniz snuffed out this single anomaly by denying that the soul acts on the body; rather in an original miracle worked at the creation of the universe, God created a pre-established harmony that caused the mechanical movements of the body to *conform* to the will of the soul, without being *caused* by it.

In short, only God's consciousness ever influenced natural events, and even this influence was confined to the origins of the universe. Once it was created, the cosmos was self-maintaining and functioned purely by means of unconscious, material events.

In the mechanical philosophy that dominated the rise of modern science, not only was nature devoid of consciousness, it was objectified to the point that it was divorced from perceptual experience altogether. The material objects that made up the world had certain primary qualities, such as size, shape and velocity; but they were inherently devoid of all secondary properties, such as color, smell, and taste, which were relative to perception. Thus, conscious experience was effectively removed from nature and, therefore, from the scope of natural science.

Over the following centuries, words that previously referred to constituents of human experience were defined in purely objectivist terms. Sound became fluctuations in an objective medium such as air; smell became molecules adrift in the atmosphere; light became a form of electromagnetic energy; and color became specific frequencies of that energy. Natural science was solely concerned with these phenomena as they were thought to occur independently in nature; and little scientific attention was focused on the manner in which such objective entities related to their corresponding events in conscious human experience.

With the twentieth century, the human mind and experience had been so long ignored by scientists that it was possible for advocates of behaviorism to deny that subjective experiences of awareness, perception, desires or feelings existed at all. But by the mid-century, the limitations of this approach for understanding human behavior were becoming increasingly apparent; and this opened the way for the development of a new field of cognitive sciences, which was to include philosophy, psychology, artificial intelligence, linguistics, anthropology, and neuroscience. A new science did indeed come

into existence, but under the domination of cognitivism, many of the old assumptions dating from the sixteenth century remained unquestioned.

Cognitivism

The foundations for cognitivism were already being laid in the late 1940s, as the limitations of the behaviorist agenda were becoming increasingly apparent to neuroscientists and psychologists alike. Rather than behavior being consequent upon environmental promptings, central brain processes were now thought to proceed and dictate the ways in which an organism caries out complex behavior. While neurobehaviorism had assumed that the nervous system is in a state of inactivity most of the time, and that isolated reflexes are activated only when specific forms of stimulation occur, this view was undermined by the recognition that the brain is a dynamic, constantly active system. Thus, rather than the organization of behavior being imposed from outside, scientists came to believe that it emanated from within the organism.[1]

One of the important components of the new cognitive science that was emerging around the turn of the century was cognitive psychology. In the late 1950s, cognitive psychology grew out of information processing. Just as such perceptual terms as color and sound had earlier been objectified, this movement rapidly redefined most cognitive terms so that they too became removed from everyday conscious experience. Information now became a non-conscious "decision" between equally plausible events, and this notion enabled scientists to focus on the efficacy of any communication of messages via any mechanism. Thus, the processing of such information could be considered apart from any particular embodiment.

N. Wiener commented on this new notion: "Information is information, not matter or energy. No materialism which does not admit this can survive at the present day."[2] But this assertion immediately raises the question: How does matter or energy *become* information, or how does information become *represented, encoded* or *located* in matter or energy? John Anderson points out that information processing is the dominant viewpoint in cognitive psychology,[3] and he claims that information "is represented in

[1] Howard Gardner, *The Mind's New Science* (New York: Basic Books, 1985), 13.
[2] N. Wiener, *Cybernetics, or Control and Communication in the Animal and the Machine*, 2nd ed. (Cambridge, Mass: MIT Press, 1961; 1st ed. 1948), 132.
[3] John Anderson, *Cognitive Psychology and Its Implications* (New York: W. H. Freeman & Co., 1990), 9.

terms of continuously varying electrochemical activity of neurons."[4] He seems to agree with the above statement by Wiener when he writes, "Perception begins with energy, such as light or sound, from the external environment. Receptors, such as those on the retina, transform this energy to neural information."[5] But nowhere does he explain the manner in which energy is transformed into information. Varela, et. al. explain the reason for this oversight: "we have no idea how the symbolic expressions supposed by the cognitivist to be encoded in the brain could get their meaning."[6]

The dominant view of modern cognitive psychology, and cognitive science as a whole, is known as cognitivism, which takes as its guiding metaphor the digital computer, which, like the brain, is said to perform computations on symbols. The computer thus becomes the mechanical model of the mind; and thought, or cognition, is identified with physical, symbolic computations. From this perspective, cognitive science becomes the study of such cognitive, physical symbol systems; and the field of artificial intelligence is a literal construal of this cognitivist hypothesis.

The textbook presentation of cognitive psychology as put forth by Anderson adopts this perspective as a matter of course. In his presentation of the neural basis of cognition, cognitive terms are uniformly objectified, as illustrated by the following examples: "Cognition is achieved by patterns of neural activation in large sets of neurons," and it "resides in patterns of the primitive elements of computers."[7] While he appears to be in no doubt in his assertion that "the brain encodes cognition in neural patterns,"[8] he acknowledges that no one knows *how* this occurs. According to Anderson, "Low level cells in the visual system detect simple patterns of spots of light and darkness in the visual field"[9]; and he would certainly agree with Daniel Dennett's comment that such subsystems "are deemed to be unproblematic *non*conscious bits of organic machinery, as utterly lacking in point of view or inner life as a kidney or kneecap."[10] Anderson offers no justification, or explanation, for asserting that such cells *detect*, rather than simply *electrochemically react to*, visually related physical stimuli.

---

[4] Ibid. 18.
[5] Ibid. 83.
[6] Francisco Varela, Evans Thompson, & Eleanor Rosch, *The Embodied Mind: Cognitive science and human experience* (Cambridge, MA: MIT Press, 1991), 3-8.
[7] *Cognitive Psychology and Its Implications*, 18 & 24.
[8] Ibid. 24.
[9] Ibid. 19.
[10] Douglas R. Hofstadter & Daniel Dennett, eds., *The Mind's Eye*: *Fantasies and Reflections on Self and Soul* (New York: Basic Books, 1981), 12.

Thus, cognitivism postulates not only unconscious cognitive processes, but ones of which *we cannot be aware.* This hypothesis that cognition can proceed without consciousness is based on the assumption that there is no essential or necessary connection between the computational mind and consciousness. Anderson extends this principle beyond cognition to emotions when he writes that "computer systems...have been shown to be capable of...displaying frustration," and this "feeling of frustration" occurs in "large patterns of bit changes."[11] He offers no evidence for the presence of this emotion, nor does he offer any insight into the manner in which feelings can become embedded in patterns of bit changes.

There seems to be a good deal of mystery surrounding the questions of how cognition and even emotions are achieved by the components of the brain and the computer. Anderson attempts to dispel this qualm by answering, "It does not appear that there is anything magical about human intelligence or anything that is incapable of being modeled on a computer,"[12] but he offers no justification for divorcing cognitive and affective terms from conscious experience and designating them upon nonconscious, material objects and processes. Contrary to his claim, there is in fact a facet of human intelligence that does not appear to have been modeled on a computer, and that is consciousness. Without addressing this issue, we are poorly equipped to answer the question: are patterns of neural activation merely necessary for conscious cognition to be achieved, or are they sufficient?

Howard Gardner writes that one of the principal features of cognitive science is the deliberate decision to de-emphasize certain factors that may be important for cognitive functioning, but whose inclusion would unnecessarily complicate the cognitive scientific enterprise. It seems clear that the de-emphasized features are those that do not conform to the computer model, which, he claims is "central to any understanding of the human mind."[13]

Ray Jackendoff has identified this limitation when he writes of two sorts of mind: the computational mind, which is readily modeled on a computer, and the phenomenological mind, which is the mind that we humans actually experience. By postulating a computational mind that is inaccessible to consciousness, he writes, cognitivism "offers no explication of what a conscious experience is."[14] His answer is

---

[11] *Cognitive Psychology and Its Implications*, 24.
[12] Ibid. 3.
[13] *The Mind's New Science*, 6.
[14]Ray Jackendoff, *Consciousness and the Computational Mind* (Cambridge, MA: MIT Press, 1987), 20.

that conscious awareness is "an *externalization* or *projection* of some subset of elements of the computational mind"[15]; but he does not claim to know which elements "project" or "support" awareness. Moreover, (perhaps unconsciously) in keeping with the early principles of mechanical philosophy, he claims that the distinctions present in the phenomenological (i.e., experienced) mind are not *made* by that mind, but are projected into it by the computational mind. Thus, consciousness remains without causal efficacy in nature; and since it cannot have any effects, it, in his words, "is not good for anything."[16]

Simply put, in this attempt to introduce the conscious mind into the discipline of cognitive science, it is attributed to an unexplained projective capability of unknown elements in the brain; and once projected, it purportedly produces no effects and is therefore good for nothing. Such a theory seems very short on explanatory power.

Connectionism

Just as cognitivism developed out of an increasing awareness of the limitations of behaviorism, so has the subsequent emergence of connectionism resulted from perceived deficiencies in the cognitivist view. Two of these deficiencies are noted by Francisco Varela, et. al.: "the most ordinary tasks, performed even by tiny insects, are done faster than is possible when attempted with a computational strategy of the type proposed in the cognitivist orthodoxy. Similarly, the resiliency of the brain to damage, or the flexibility of biological cognition to new environments without compromising all of its competence...is...nowhere to be seen in the computational paradigm as such."[17]

In the connectionist view, brains "operate on the basis of massive interconnections, in a distributed form, so that the actual connections among ensembles of neurons change as a result of experience. In brief, these ensembles present us with a self-organizing capacity that is nowhere to be found in the paradigm for symbol manipulation."[18] The notion of representation is maintained in connectionism, but it has come to mean the correspondence between an emergent global state and properties of the world. Thus, unlike cognitivism, representation is not thought to be a function of particular symbols. Moreover, connectionism refutes the cognitivist view of cognition as the representation of a pre-given world by a pre-given mind; and it replaces this notion

[15] Ibid. 23.
[16] Ibid. 26.
[17] *The Embodied Mind*, 5-4.
[18] Ibid. 5-2.

with the assertion that cognition *enacts* a world as a domain of distinctions that is inseparable from the structure embodied by the cognitive system.

Varela et. al. illustrate this view with an analysis of color perception. When we actually measure the light reflected in the world around us, they write, we "discover that there simply is no one-to-one relationship between light flux at various wavelengths and the colors we perceive things to have."[19] Since the properties that specify what perceived colors are have no non-experiential physical counterparts, we cannot account for our experience of color as an attribute of things in the world by appealing simply to the intensity and wavelength composition of the light reflected from an area. This has direct implications for visually perceived objects in general, since it is color contrast that visually forms those objects.

According to these authors, the study of color illustrates that colors are not "out there" in the independent physical world, nor are they "in here" in the subjective mind. Neither the world we cognize nor the mind that cognizes it is pre-given; rather, both the world and the perceiver specify each other and are enacted through their history of structural coupling.

To sum up, connectionism views cognition as the emergence of global states in a network of simple components, which operates through local rules for individual operation and rules for changes in the connectivity between the elements. Meaning, instead of being located in particular symbols, is thought to be a function of the global state of the system. Instead of residing in specific constituents of the brain *per se*, it is said to reside in complex patterns of activity that emerge from the interactions of many such constituents.

Towards an Understanding of Consciousness

Connectionism, as presented by Varela et. al., regards conscious experience as a global emergent property of the connection between an organism and its environment. These authors point out that emergent properties commonly occur in densely connected aggregates of any kind,[20] but they fail to explain how the phenomenological, or experienced, mind emerges in some living organisms but not, presumably, in non-organic compounds. Another important issue that they fail to address is the set of specific properties of the phenomenological mind that distinguish it from unconscious matter and

[19] Ibid. 8-13.
[20] Ibid. 5-8.

energy. Their assertion that the mind and the world it cognizes specify each other tends to blur any distinctions that may remain between the two. Even in the participatory universe envisioned by these authors, it must still be possible to speak of the ways in which the body and the environment influence the phenomenological mind; and an earlier question can be posed once again: To what extent, and in what ways does this mind exert a causal influence upon the physical world?

For the cognitivist, meaning is located in specific constituents and processes of the brain, whereas for the connectionist it resides in complex patterns of activity that emerge from the interactions of many neural and environmental constituents. But neither theory explains the *nature* of such meaning, or information, nor do they explain how it got there. It does not appear that either view has fully come to terms with the split between objectivist and experiential uses of such terms as color, light, vision, cognition, information, and meaning. Nor does either hypothesis satisfactorily account for the causal interactions between objective and experiential events.

Anderson's cognitivist discussion of visual perception is worth reviewing with this point in mind. "Light," he writes, "is converted into neural energy by a photochemical process," and low-level cells in the visual system "detect simple patterns of spots of light and darkness in the visual field."[21] Objective light is thought to consist of quantized electromagnetic energy of various frequencies, traveling through space at 186,000 miles per second. It is quite feasible for such energy to be converted into objective neural energy, *but there is no account of any process by which such objective energy becomes transformed into experiential patterns of lightness and darkness in a visual field.* It is perfectly feasible for non-conscious, low-level cells to *react* to impulses of energy, but we are given no reason to believe that they *detect* experiential patterns of lightness or darkness in a perceived visual field. Certainly such an experiential visual field does not travel through space at the speed of light, but we are given no explanation for its occurrence at all. And this omission is concealed by the *objectivist* use of the cognitive term *detect* with respect to an *experiential* visual field. The mind/body problem has not been solved by this ploy; *it has been camouflaged by terminological sleight of hand.*

This point has direct bearing on the endeavor to design computer models to simulate the information processing that takes place in the visual system. According to Anderson, the goal of such research is "to get these computer programs to see in a visual

[21] *Cognitive Psychology and Its Implications*, 33 & 19.

scene what a human sees."[22] Humans experientially see a visual world in dependence upon unconscious interactions between the environment and the nervous system; and according to the research cited by Varela, et. al., there is no one-to-one correspondence between the frequencies and intensities of objective light and the experiential patterns of light that make up our visual world. If Anderson is suggesting that the computer programs he has in mind are designed to *consciously experience* the visual world perceived by humans, he must identify which components of these programs produce consciousness, and he should explain the mechanism for this occurrence. If, on the other hand, such computers are designed to unconsciously react to objective light, then there must be a one-to-one correspondence between their reactions and the objective features of the incoming electromagnetic energy. But in this case, according to the evidence cited by Varela, et. al., it would follow that those mechanical devices are guaranteed not to "see in a visual scene what a human sees." Indeed, we have no reason to believe that they experientially *see* anything.

In the connectionist view propounded by Varela et. al., the experienced world arises in connection with conscious experience; and that world does not exist in a pre-given, objective manner. What is the relation between this experienced world and the objective world posited by natural science? The electromagnetic energy that makes up objective light, the vibrations in the atmosphere that constitute sound, and the other objective correlates of perceptual experience are commonly thought to exist in objective space-time in the outer world. Where does the experiential world posited by this connectionist view exist?

Our authors deny that it exists either in the independent, objective world, or that it exists in the mind. The former assertion is easily defended. For example, the visual patterns we see are contingent upon our visual faculties, and in the absence of those, or similar faculties, visual objects do not exist in nature at all. Thus, unless our visual faculties are endowed with a supernatural ability to project visual images out into objective space, there seems to be no reason to locate them in the objective world. Following this same reasoning, we must conclude that the visual objects we perceive do not exist out there in the real, external world; and, by the same logic, this must be true of the entire experienced world.

The notion that the objects that make up our perceived world exist in our heads is physically implausible, to say the least. No one's head is that big; and besides, the head is already filled with other things. But if the perceived world does not exist out there or in

[22] Ibid. 37.

here, there seems to be no place for it at all in objective space. But how can the world we presently experience be nowhere? And, assuming that the objectified world of science, with its matter and energy, does exist in objective space-time, how could this real physical world causally contribute to our experienced world, which apparently has no objective location, and possibly no objective spatial dimension?

Cognitive psychology has drawn our attention to the existence of mental imagery, and Anderson writes that, "when subjects are scanning a mental array, they are scanning a representation that is analogous to [a] physical array."[23] Thus, a mental image is "an abstract analog of a spatial structure," and certain data might seem to indicate that subjects rotate mental objects in a three-dimensional space "within their heads."[24] Anderson hastens to add that "subjects are not actually rotating an object in their heads,"[25] but he does not explain where mental objects are rotated. If not in the head, it is even less likely that such objects exist outside of the head; and this leaves them nowhere at all.

We may well wonder, at this point, whether the non-locality of the perceived world is the same or different from the non-locality of the imaginal world. There is evidence that the perceived and imaginal worlds are causally related. For example, in biofeedback research it has been found that when subjects imagine warmth in their hands, their fingers warm up. Does this suggest that a mental image, say, of a flame that is imagined in one's hand is actually present in the experiential space occupied by the hand (bearing in mind that we have been unable to locate the phenomenological hand, or any other experienced object in objective space)? If this were the case, it would seem plausible that an object imagined outside of the body might exist in the experienced space outside the body; and there seems no reason in principle to reject the possibility that such an imagined object might influence the experienced objects where it is imagined.

Such questions are not apparently raised in either the cognitivist or the connectionist views, for neither has clearly confronted the difference between the objective world posited by scientists and the experiential world that we inhabit. Nor does either view compellingly address the causal efficacy of the experienced mind in nature. Varela et. al. do comment that "A rule for the constitution of the brain is that if a region (nucleus, layer) *A* connects to *B*, then *B* connects reciprocally back to *A*."[26] This corresponds to a similar principle in natural science as a whole, but the notion that the

---

[23] Ibid. 96.
[24] Ibid. 98 & 93.
[25] Ibid. 93.
[26] *The Embodied Mind*, 5-11.

computational mind and the phenomenological mind may bear a reciprocal, causal relationship has been largely ignored until very recently. Current scientific concern with this issue appears under the label *psychoneuroimmunology*, and there is increasing evidence that research in this area may yield substantial results in the fields of physical and mental health.

The possibility that the experienced mind may have a reciprocal, causal relation with the experienced world beyond one's own body would be the next area to be researched following this line of inquiry; but such research has barely begun.

As mentioned previously, the origins and nature of meaning, information, and representation remain unsolved in both the cognitivist and the connectionist theories. In seeking an answer to these questions, let us take as an example the word *pot*. If we do not consciously attribute meaning to this sound, the sound, of itself, is just a noise that means nothing, represents nothing, and conveys no information. In this regard it is no different than the sound of thunder or the wind. Even when we have attributed meaning to this sound, it remains only a meaningless noise for non-English-speaking people (assuming they haven't attributed their own meaning to this sound). Thus, the conscious designation of meaning upon a sound is necessary for it to become a word. Without such a designation, nothing can be said to have meaning, and nothing represents or conveys information about anything else. The factor of consciousness is indispensable, yet it appears to have no role in information theory, and it is largely overlooked in cognitive psychology and cognitive science as a whole.

Meaning originates from conscious designation, and its nature in any specific case is relative to the consciousnesses that have adopted the given designation. Thus, the English language is meaningful for those who have learned the meaning of its words and grammar, but it is meaningless for those who have not.

Recall Anderson's cognitivist assertion that information is represented in terms of continuously varying electrochemical activity of neurons. Just as English-speakers can meaningfully assert that information is represented in the sound *pot*, so can cognitivists meaningfully make the above assertion about electrochemical activity. But apart from the respective designations by English-speakers and cognitivists, neither sound nor electrochemical activity inherently represents any information whatever. This statement does not preclude the possibility that certain types of electrochemical activity may be *physically necessary* for certain types of conscious, cognitive events to take place. The point is that all such electrochemical events can be described without the use of such cognitive terms as *information, meaning, represent, detect, see, remember,* and *judge;* and with this demystified shift of terminology, the physical basis of the

phenomenological mind can be explained just as well as it is with the use of those terms...and perhaps with less confusion.

Cognitive science has ignored the nature of consciousness in much the same manner that physical science long ignored the role of the observer and experimenter in the physical world. Research in quantum mechanics has resulted in the conclusion that the observer can take any totality apart from one including himself in the act of performing an experiment, but he (or at least a part of himself) must always lie outside the system.[27] A similar conclusion has been drawn in the field of linguistics, as Hilary Putnam comments: "I can generalize over as large a totality of languages as I want (excluding my own language), but the language in which I do my own generalizing must always lie outside the totality over which I generalize...The 'God's-Eye View'—the view from which absolutely all languages are equally part of the totality being scrutinized—is forever inaccessible."[28]

Consciousness lingers outside the parameters of contemporary cognitive psychology, not as a disembodied spirit, but as an indispensably active participant in the scientific enterprise and in human experience as a whole. There is no question that consciousness plays a crucial, causal role in the creation of science, just as it does in the occurrence of the humanly perceived world and the imaginal world. It may well be that the objectified world of natural science is also produced only in relation to the consciousnesses that conceive it. Be that as it may, the causal efficacy of consciousness in nature appears to be a frontier for scientific research that will strike at the root of some of the most fundamental assumptions that have underlaid modern science since its origins.

---

[27] Hilary Putnam, *Realism with a Human Face*, ed. James Conant (Cambridge, Mass.: Harvard University Press, 1990) p. 17.
[28] Ibid.