

CONSCIOUSNESS AND THE MACHINE

by Eugene G. d'Aquili and Andrew B. Newberg

Abstract. We consider only the relationship of consciousness to physical reality, whether physical reality is interpreted as the brain, artificial intelligence, or the universe as a whole. The difficulties with starting the analysis with physical reality on the one hand and with consciousness on the other are delineated. We consider how one may derive from the other. Concepts of universal or pure consciousness versus local or ego consciousness are explored with the possibility that consciousness may be physically creative. We examine whether artificial intelligence can possess consciousness as an extension of the interrelationship between consciousness and the brain or material reality.

Keywords: artificial intelligence; brain; consciousness; material reality; neuroepistemology; subjective awareness.

I. INTRODUCTION

Originally, this paper was to consider only the question of whether artificial intelligence could possibly be conscious, either now or as a result of future technological advances. The issue is fascinating, especially since the relationship of consciousness to the brain is anything but clear. Therefore, to understand the problems of consciousness and artificial intelligence, it is first necessary to understand the problems that consciousness poses in general and particularly how it may relate to any physical reality, whether this be the brain or an elaborate computer. To an adult human being with a normally functioning brain, reality, at first pass, seems to be composed of two vividly real categories: the conscious self, and external reality comprising things

Eugene G. d'Aquili, Clinical Associate Professor of Psychiatry at the University of Pennsylvania Medical School, may be reached at Limina, 2400 Chestnut Street, Suite 1503, Philadelphia, PA 19103. Andrew B. Newberg, a resident in internal medicine at the Graduate Hospital of the University of Pennsylvania, also may be reached at Limina, 2400 Chestnut Street, Suite 1503, Philadelphia, PA 19103. D'Aquili presented a version of this paper at the Forty-Second Annual Star Island Conference of the Institute on Religion in an Age of Science (IRAS), "Life in the Universe," at Star Island, New Hampshire, 29 July-5 August 1995.

[*Zygon*, vol. 31, no. 2 (June 1996).]

© 1996 by the Joint Publication Board of *Zygon*. ISSN 0591-2385

that appear to have an inherent reality separate from the conscious self. The things in external reality also appear to be represented in, or known by, the conscious self. Thus, the classical philosophical problem of subjectivity versus objectivity is a problem only because the brain, under ordinary conditions, insists on processing reality in this manner. To the naive observer, there is an absolutely certain sense that there is a reality external to the self that appears to be heavy and substantive; this is often termed *matter* or *material reality*. The naive observer also has the absolutely certain sense of a conscious self that seems to have a light, changeable, and ethereal quality; it is often termed *mind*, *spirit*, or sometimes *soul*. The naive terminology is anything but exact.

These two senses of reality are so vivid and appear so real that early philosophy did not seriously question the fundamental nature of this duality. For the first thousand years of its existence in the West, philosophy began its work by concentrating primarily on the substantiality of reality. This was the heyday of ontology. However, beginning with Descartes, followed by the radical empiricism of Bishop Berkeley, among the British empiricists, and with Immanuel Kant, on the Continent, there was a shift in emphasis, to mind as the philosophical starting point, and to how we can know external reality, or anything at all for that matter. Since the seventeenth century, therefore, we have seen the heyday of epistemology with its emphasis on how we know and, in its extreme manifestation, the assertion that all reality is mind. At first, modern science waded in on the side of the old ontology, naively assuming the existence of external reality as represented in consciousness. Since the turn of the century, and particularly since the development of quantum theory, science finds itself caught between ontology and epistemology, with old certainties vanishing like smoke. Bertrand Russell expressed the state of this confusion manifested in the conversation around his childhood dinner table. His parents apparently were given to discussing the nature of reality. The joke was, "Is reality mind? No matter! Is reality matter? Never mind!"

As just noted, the problem of the relationship of subjective awareness to external material reality began to achieve a modern focus with Descartes's clear and unabashed dualism. Descartes saw the mind as a subjective awareness that contained ideas corresponding, or sometimes not corresponding, to what was in the external world. For Descartes, the mind ordinarily represented the world in a one-to-one correspondence except for the occasional glitches which generated error (Descartes 1911). This view of the mind as representational of the external world reached its apogee in the work of Franz Brentano. According to Brentano, all states of awareness are *of* or *about* something. For Brentano, mental states must necessarily have "reference to a content" or "direc-

tion toward an object," which characteristic of mental states he called *intentionality* (Brentano 1973, 88). This vivid directedness or intentionality was, for Brentano, the defining characteristic of the mind.

Edmund Husserl, often called the father of phenomenology, was one of Brentano's students who began his work trying to develop a specific procedure for examining the structure of intentionality, which, of course, was the structure of experience itself, without making any reference whatsoever to a factual empirical world and especially without any assumption of its actual existence. This rigid formal procedure he called *epoche* or "bracketing," for it required that one "bracket," or suspend belief in, one's ordinary judgments about the relation between experience and the world "out there." Husserl maintained that these ordinary judgments that must be temporarily suspended or bracketed represent a "natural attitude" that, when it is raised to the level of a philosophical school, is called *naive realism*. By bracketing what he came to see as the "hypothesis of the natural attitude" Husserl attempted to study the intentional contents of the mind purely internally, that is, without tracing them back to what they *seemed* to refer to in the external world. By this approach, he claimed to present a new domain of knowledge that was absolutely prior to any empirical science. Starting with pure experience, and eschewing all assumptions implicit or explicit about the nature of reality, Husserl embarked upon a sort of philosophical introspection that he called *Wesensschau* or "intuition of essences." By this process, Husserl attempted to reduce experience to essential structures and then demonstrate how our human world was generated from them (Husserl 1931). We can now see how far Husserl's approach to subjective awareness eventually led him from Descartes's and Brentano's mental representationalism (Husserl 1970). Husserl's rigorous phenomenological approach had put an independent isomorphic external world on very shaky ground indeed.

Expanding upon the work of Husserl, Maurice Merleau-Ponty recognized the enormity of the problem for both science and philosophy of trying meaningfully to relate subjective awareness to the vivid sense of external reality or to the sense of world. In his *Phenomenology of Perception*, Merleau-Ponty (1962, x-xi) wrote: "The world is inseparable from the subject, but from a subject which is nothing but a project of the world, and the subject is inseparable from the world, but from a world which the subject itself projects." Merleau-Ponty clearly recognized the need for a bridge between self and world, between the apparent inner and the apparent outer. Unfortunately, his astuteness in defining the problem was not matched by his ability to solve it.

If we add to all of this rigorously derived confusion the current ferment among cognitive scientists of various persuasions, attacking

each other and rushing off to write books claiming to explain consciousness based on the most recent blips on an oscilloscope or upon the most recent imaging studies of the brain, we can begin to understand the chaos surrounding the relationship of subjective awareness to external material reality. It is probably the most important general scientific and philosophical problem of our time. And, we must point out, the possibility of the consciousness of artificial intelligence is inextricably interwoven with these issues.

An old philosophy professor of our acquaintance maintained that if our philosophy begins with the reality of external existence, it can never adequately explain the development of consciousness. And if our philosophy begins with consciousness, it can never adequately explain the development of external reality. Let us consider these two approaches separately to see if we can develop a reasonably integrated approach, first to the consciousness/brain problem and then to the consciousness/artificial intelligence problem.

II. IF EXTERNAL REALITY IS PRIMARY

If external reality is accepted as primary, the question which we must answer is, How is consciousness generated by the brain and nervous system, and how can this be understood in an evolutionary context?

We must realize that neuropsychology up to the present, and parallel to Brentano's philosophy, always has understood consciousness to refer to consciousness of something. That pure consciousness, devoid of content, might exist has generally not even been entertained as a problem. Therefore, obviously, there has been little attempt at understanding its physical basis. We will return to the issue of pure consciousness below. First, let us consider the basic and classical neuropsychological problem of how consciousness of anything is possible. Here we are using *consciousness* in its very simplest sense of awareness. We do not mean consciousness of Self, or how the Self comes to be conscious. In other words, we are not referring to the reflexive self-consciousness that human beings are capable of. In this context, we are simply referring to consciousness as subjective awareness, whether in lower animals or in human beings.

To this point, we have been using the words *consciousness* and *awareness* interchangeably in a naive sense. We must now define these concepts more carefully and use them more precisely:

We define *subjective awareness* as any and all mental content that inheres in a subject, excepting only a reified sense of Self.

We define *consciousness* as any and all mental elements which inhere in a subject, one of which elements is a reified sense of Self.

Strictly speaking, consciousness involves the generation of a Self as an element in subjective awareness. We have proposed that the mind/brain becomes aware of a certain set of sensory input that ultimately arises from the body or from the body's interaction with the external world. In other words, the mind/brain perceives its multiple activities and organizes them into a reified category that we call the Self. Considered evolutionarily, such a process becomes possible only with the evolution of the inferior parietal lobule and its interconnections with various sensory association areas. These structures are known to underlie the reification of classes of objects generating abstract categories. If this is so, then the neuroanatomical requirements of "selfhood" must restrict the clear sense of Self to higher primates, and especially to *Homo*. There is, in fact, good evidence that this is so. For example, only higher primates respond to their image in a mirror as if it were a representation of themselves. All other animals apparently perceive another beast.

Simply put, subjective awareness is consciousness without a clear reified Self, and consciousness is subjective awareness with a reified Self.

Finally, the inferior parietal lobule and interconnected sensory association areas can operate on, and reify, the Self perceiving the Self, generating what has been called reflexive consciousness. It is generally thought that clear reflexive consciousness is a property only of *Homo sapiens*. However, this is still an open question; some anthropoid apes may possess it.

Having more carefully defined out terms, we must note that we are more concerned in this paper with the implications of simple subjective awareness, common, to one degree or another, to all sentient creatures, than we are with the evolution of consciousness of "Self" and of reflexive consciousness.

Perhaps the biggest problem currently faced by neurophysiologists and neuropsychologists is how recognition or awareness of sensory input comes about. This has come to be known as the *binding problem*. It is known that the brain breaks down sensory input into many constituent parts. These constituent parts are analyzed and physically stored in different parts of the brain. Within each sensory modality, the specific identifying elements or "recognition features" are stored in physically different locations within the brain's association areas for that modality. For recognition of a sensory input, it is necessary to somehow bring together at least the essential recognition features both within a given sensory modality and then across modalities. One can see the magnitude of the problem when one realizes that any given sensory input is broken down according to the various sensory components: visual, auditory, tactile, and possibly olfactory and gustatory as well. Each of these elements of the sensory input is then divided into many recognition features that are physically stored separately. Most

neuroscientists agree that the binding of these disparate elements to match new sensory input, generating recognition, is a temporal and not a spatial phenomenon. In other words, the various recognition features stored physically apart are probably not reassembled in one place in the brain, like a mosaic, to form a picture, which is then compared with input, thus generating recognition. Rather, it seems that all of the relevant features are somehow activated where they are stored at precisely the same time. It is this temporal binding of recognition features that apparently underlies conscious awareness and recognition.

R. Llinas (1988), at New York University, has intriguing evidence regarding how this is possible. Recent extension of his earlier work indicates that the initial breakdown of recognition features either has already occurred prior to the stimuli reaching the thalamus or else occurs at the thalamic level. The cells in the thalamus that encode each recognition feature send pulses of a certain frequency that generate corresponding pulses in the neurons storing these recognition features throughout the brain. The simultaneous pulses of the same frequency in all the relevant storage areas, even across sensory modalities, somehow result in the binding of these features resulting in recognition of the incoming material. According to Llinas, just imagining a scene may involve the reverse phenomenon: multiple bits of stored information all begin to pulse at the same frequency at the same time, creating an activation of corresponding cells in the thalamus. Thus, awareness and recognition of incoming sensory input, or imagining objects in a scene, all result from a sort of thalamic-cortical dialogue. This model, which would solve much of the binding problem, allowing awareness of our external environment or even of imagined entities, still has many problems and raises many questions. Almost everyone agrees that the amygdalae and hippocampi are somehow involved in the generation of awareness, both via their connections with the thalamus and via their connections with sensory association areas as well as with other cortical regions. The precise role of the amygdalae and hippocampi is not yet clear in Llinas's model.

Whatever may be the ultimate mechanism of binding that underlies conscious recognition and imagination, it is becoming clearer that the association areas involved with each sensory modality seem to be somehow responsible for subjective awareness in that modality. The evidence is particularly strong with vision. The phenomenon of blindsight, first described before the turn of the century, is a rare but well-documented pathological condition. In this disorder, the primary visual area (the calcarine cortex) is left intact, but there is destruction of most of the visual association areas. In this condition, the patient claims to be totally blind. He or she can, in fact, see nothing, at least consciously.

However, such patients can negotiate through rooms of furniture, through doors, up and down staircases, and even on busy streets without ever bumping into anything or taking unusual chances. In other words, they behave as if they were sighted without having any conscious visual awareness of the world around them. There is some evidence that a similar condition may obtain with other sensory modalities when their association areas are destroyed but their primary cortical areas are intact. If further evidence bears this out, especially in sensory modalities other than vision, it is reasonable to assume that subjective awareness arose with the evolution of secondary sensory association areas.

The binding problem and the specific function of the sensory association areas are two major issues that neuropsychologists are currently investigating to obtain an understanding of subjective awareness either of the external world or of imagined gestalts. All of this discussion refers to the mechanisms underlying *awareness of something*.

Over the years, we have become interested in understanding *Pure Consciousness*, i.e., consciousness devoid of content, sometimes described as a clear and vivid consciousness of nothing, or perhaps of everything at the same time. In previous works (d'Aquili and Newberg 1993a, 1993b), we have described in detail a state that we have called Absolute Unitary Being (AUB). AUB is described in the mystical literature of all of the world's great religions, and it has been attested to by modern secular mystics. Erwin Schrödinger, the father of quantum theory, and in theory the owner of a famous cat, reflected on the significance of his experience of what seems to be AUB with these words: "The only possible alternative (to the plurality of souls hypothesis) is simply to keep to the immediate experience that consciousness (i.e., Mind) is a singular of which the plural is unknown; that there *is* only one thing and that what seems to be a plurality is merely a series of different aspects of this one thing, produced by a deception; the same illusion is produced in a gallery of mirrors, and in the same way Gaurisankar and Mount Everest turned out to be the same peak seen from different valleys" (Wilber 1975). In another place, Schrödinger states: "Inconceivable as it seems to ordinary reason, you—and all other conscious beings as such—are all in all. Hence, this life of yours you are living is not merely a piece of the entire existence, but is in a certain sense the *whole*. . . . Thus, you can throw yourself flat on the ground, stretched out upon Mother Earth with a certain conviction you are one with her and she with you. You are as firmly established, as invulnerable as she, indeed a thousand times firmer and more invulnerable" (Wilber 1975). Of the modern secular mystics, in addition to Schrödinger, we can add J. Robert Oppenheimer, Niels Bohr, and a number of other theoretical physicists. Dag Hammarskjöld, an early and influential secretary general

of the United Nations, was among the modern Western secular mystics who have described the state of AUB. As we have said, AUB is described by mystics of all the world's religions. Here we will let one example suffice. Ch'an Master Huang Po wrote: "All the Buddhas and all sentient beings are nothing but One Mind, beside which nothing exists. This Mind, which is without beginning is unborn and indestructible. It is not green or yellow, and has neither form nor appearance, it does not belong to the categories of things which exist or do not exist, nor can it be thought of in terms of new or old. It is neither long nor short, big nor small, for it transcends all limits, measures, names, traces, and comparisons. Only awake to the One Mind" (Wilber 1975). In AUB, there are no boundaries of discrete beings, there is no sense of the passage of time, no sense of the extension of space, and the self-other dichotomy is totally obliterated. In other words, the state consists of an absolute sense of unity without thought, without words, without sensation and even without any sense of inhering in a subject. We have proposed that total deafferentation of the posterior superior parietal lobule, especially on the right, results in this state. This area of the brain is responsible for the orientation of objects in three-dimensional space (Lynch 1980; Mountcastle 1976; Mountcastle, Motter, and Anderson 1980). If it is denied of all input as a result of mechanisms generated during profound meditation, it creates a sense of pure space. Since space has no subjective reality unless it relates things to each other, the subjective experience is one of total spacelessness or of total perfect unity. It is interesting that there is evidence that the posterior superior parietal lobule in the left hemisphere may be responsible for the self-other dichotomy (Joseph 1990). During profound meditation, we have proposed that the posterior superior parietal lobule on both sides is totally deafferented, resulting not only in the sense of absolute space, but in the obliteration of the distinction between self and other (d'Aquili and Newberg 1993a, 1993b).

Over the past year, we have studied practitioners of Tibetan Buddhism, who have meditated an average of one to three hours per day for the past fifteen years (Newberg, Alavi, et al. 1995; Newberg, Baime et al. 1995). We have studied them using a nuclear imaging technique known as single photon emission computed tomography (SPECT). Our pilot study so far strongly supports the model we have proposed. Thus, it seems that a state of Pure Consciousness can be achieved via intense meditation by deafferentation of a certain part of the parietal lobe bilaterally.

III. COMMENTARY ON BEGINNING WITH EXTERNAL REALITY

The problem with everything that we have discussed up to this point is that while these neurophysiological mechanisms are correlated with

awareness and may even be the causes of awareness, they do not explain the stuff of awareness itself. In this regard, Penrose notes:

If it were not for the puzzling aspects of consciousness that relate to the presence of "awareness" . . . which as yet seem[s] to elude physical description, we should not need to feel tempted to look beyond the standard methods of science for explanation of minds as a feature of the physical behavior of brains. . . . It may well be that in order to accommodate the mystery of the mind, we shall need a broadening of what we presently mean by "science," but I see no reason to make any *clean break* [italics ours] with those methods that have served us so extraordinarily well. (Penrose 1994, 50)

We can certainly agree with Penrose that a clean break with traditional science is neither required nor desirable. But we strongly agree as well that a broadening of what is meant by science, perhaps a total realignment toward cognitive science, is required by any systematic study of consciousness or awareness.

If one looks at the traditional Aristotelian four types of causality that were considered necessary to explain a phenomenon fully, i.e., efficient causality, material causality, formal causality, and final causality, we find that our scientific explanation of awareness satisfies only one of the four requirements, efficient causality. Efficient causality is knowledge of a phenomenon in terms of anterior sequential causes. It is what we ordinarily mean by causality in modern parlance. Material causality is knowledge of the constitutive substance of the phenomenon. Clearly, we do not have the vaguest idea of what the stuff of awareness actually is. Formal causality is knowledge of a phenomenon in the organization of its constituent parts. Awareness itself has no constituent parts. The contents of awareness are its objects and not part of what it is itself. It would seem that awareness itself is simple and hence has no formal cause. Final causality is a knowledge of things in their purpose, or, in modern terminology, in terms of their adaptive function. Although final causality as originally formulated is subject to the critique of teleology, its reformulation as teleonomy has an important function in the philosophy of science.

Although it may be counterintuitive, we intend to show that awareness itself has no adaptive function. Therefore, awareness is a unique phenomenon to analyze. We can understand its physical causes, and the evolutionary adaptability of its physical causes, but that is all. Understanding awareness in terms of its physical causes, or understanding the multiple physical causes corresponding to the many events occurring in awareness, is called *psychoneural correlation*. Some scholars, desperate to bridge the gap between the subjective and the objective, have proposed *psychoneural identity*. This is an attempt to bulldoze through the problem by maintaining that the neural events themselves are awareness. This

position is not that the neural events cause awareness, or are correlated with conscious phenomena, but that they are the very thing itself. One does not have to be an epistemological genius to see the problem here. It is tantamount to saying that the machinery of an automobile is the movement of the automobile itself, or that the structure of a computer is the solution it generates to a problem. Obviously, we are not going to be let off the hook so easily.

This problem of subjective awareness has frequently driven otherwise careful scholars in the opposite direction, either flirting with or outright embracing mentalism. Even such a brilliant mathematical theoretician as Kurt Gödel rejected Turing's contention that "there is no mind separate from matter" and called this view a "prejudice of our time." Partially deriving his position from the implications of his famous theorem, Gödel maintained that although the *physical* brain must itself behave computationally, the mind is something beyond the brain. In his view, the mind is not constrained to behave according to the computational laws that he believed must control the brain's behavior (Gödel 1990, 297). In this way, Gödel's view is as extreme as the view of those who maintain psychoneural identity in the opposite direction. All this is by way of emphasizing the extremely problematic nature of the question of consciousness or awareness.

This brings us to the biggest problem of all when we begin our analysis with the primacy of external reality. The problem is, Why should subjective awareness exist at all? If every change in awareness, every change in the contents of awareness, and even if the generation of Pure Awareness itself, are all caused by physical (i.e., neural) events, as we believe they are, then why should awareness exist? There is no reason why biologically evolved robots with no subjective awareness could not have produced the entire social universe that we know, with every product of our individual endeavors, every product of our social interactions, and, in short, every psychological or cultural product, from science through art and religion. The central nervous system is an electrical input/output system of immense complexity. However, it is no more than that, or so it would appear. No matter what degree of complexity the nervous system has attained or will attain in the future, this complexity never implies in itself the existence of subjective awareness. It might produce the appearance of subjective awareness to an external observer, but there is no reason why subjective awareness should in fact exist. Actually, there is no reasonable hypothesis to explain how subjective awareness could arise out of any electrical input/output system. The material nature of the causes of awareness and awareness itself, we maintain, are incommensurables, although obviously awareness depends on its neurophysiological substrate. Again, all

this is true only if we begin the analysis with external material reality.

One often hears it said that subjective awareness had to evolve because input becomes so complex in higher animals that awareness is required to process it. But if the material, physical, and neural processing of the brain is not adequate for the job, then are we to believe that subjective awareness comes into being and somehow lifts off its neuroanatomical base, as it were, to perform the required complex analyses and then presumably settles down again on its physical base? If such a circumstance were to occur, if awareness for even a moment were free of its neurophysiological base, then we would most certainly have the ghost in the machine. In such a case, we would have demonstrated a substantial soul separate from neurophysiological functioning.

However, if awareness never lifts off its neurophysiological base—as indeed we cannot believe it does—then what is the purpose of awareness? All the physical mechanisms underlying awareness are operating to analyze reality and to respond to it. It would seem that subjective awareness is epiphenomenal. Again, why should such an epiphenomenal reality exist at all? We do not claim to have the answer to this question, but we are left with a mystery and a fundamental paradox.

It seems to us that if we start our philosophical analysis with the reality of matter and the external world, then there are fundamentally two great discontinuities in the universe. The first discontinuity is the Big Bang, or more specifically why there is something rather than nothing. This is, of course, the question that plagued Heidegger and many philosophers since. The second great discontinuity in the physical universe is the existence of subjective awareness. It simply represents an unexplainable jump from material organization to a level of reality of another order, analogous to the jump from nothing to something. Sometimes one hears it said that the evolution of life represents a discontinuity as well. We cannot agree with this since there is a smooth transition in the evolution of molecular complexity, in the development of mechanisms of replication, and with respect to the other characteristics we associate with life. But, as with the Big Bang, as with something rather than nothing, there is a fundamental discontinuity in the evolution of subjective awareness. Again, we must keep in mind that all this is true only if we assume the primacy of material reality as our philosophical starting point.

Another major problem with assuming the priority of external physical reality is the problem of isomorphism between subjective awareness and the external world. This can be summarized by the simple question, How do we know that the world as known to us corresponds in some significant way to the external world? Since we have begun our analysis by acknowledging the priority of external material reality over subjective awareness and by assuming that subjective awareness evolves from the evolution of

physical and biological systems, we can fairly safely assume that there is some degree of isomorphism between external reality and subjective awareness, or else the organism simply could not adapt to the world in which it finds itself. But what degree of isomorphism is required for an individual organism to adapt to its environment? The answer simply is not known. Perhaps only a mild to moderate degree of isomorphism is biologically required to select for an organism's perceptual apparatus. There is actually a fair amount of evidence that different species perceive the world in remarkably different ways, but always in a manner conducive to their survival. If such is the case, assuming that subjective awareness evolves from matter can lead only to an epistemological inference of adaptiveness, never to truth. If the traditional definition of truth as the *adaequatio intellectus ad rem* is taken at all seriously, then truth can be at best an approximation, perhaps a fairly weak one at that.

To this point, we have been considering the pros and cons of an analysis of the world in which external material reality is primary and subjective awareness is presumed to derive from it. However, a careful phenomenological analysis can strongly challenge this basic premise. Indeed, as Husserl implied, from the point of view of any careful conscious examiner of the world, the only thing that is certain is that all of material reality, including the laws of science and the brain itself, exists within subjective awareness. Whether it has any other substantive reality is an open question, but what is certain is that material reality exists within awareness. Furthermore, what also exists within subjective awareness is the vivid sense that the external world is substantively real and that matter is something other than consciousness. But this vivid sense, which has been called *phantasia catalyptica* by the Stoics, intentionality by some phenomenologists, and *Anwesenheit* by certain modern German philosophers, likewise exists within awareness or is an aspect of awareness. Thus, it would appear that all the vividness of the reality of the material world is at least a subset of awareness, whatever else that vividness may or may not imply. So let us see what happens to our analysis of the relationship between subjective awareness and external material reality when we give subjective awareness ontological priority.

IV. IF SUBJECTIVE AWARENESS IS PRIMARY

We might ask, What are the advantages and disadvantages of granting primacy to subjective awareness in an analysis of the relationship of subjective awareness and external material reality? The greatest advantage is that the problem of explaining the development of subjective awareness evaporates, since subjective awareness is the fundamental given matrix which permeates everything. In this case, the problem is to explain how external material reality in some sense arises out of subjective

awareness. Again, from this perspective, all of physical reality exists in present subjective awareness, including the knowing brain, all the laws of science, the compelling sense of the otherness of external material reality, the compelling sense of a past of completed events and of a future of possible ones. Since all of material reality exists *at least* in the mind of the analyzing knower, and since one would have to step outside of subjective awareness to ascertain if any reality other than subjective awareness exists (a patently impossible situation), then one is constrained to see material reality (its past and future), the laws of nature, and science itself as aspects of present subjective awareness. As disagreeable as such an epistemological position might be to those trained in Western science, it is the only possible rigorous stance unless one wishes to make a complete act of faith that the vivid sense of the otherness of external reality, which certainly exists in subjective awareness, reflects an isomorphic referent outside of subjective awareness.

From a pragmatic point of view, such an act of faith is not so terrible. We all make it almost all the time, and we use it as a basis for our actions. But if one wishes to take a rigorous phenomenological approach, it is clearly impossible to get outside of subjective awareness to determine the existence of a corresponding alternate reality. One clear advantage of approaching the problem via the priority of subjective awareness is that, in such a system, there are no discontinuities. The Big Bang itself becomes an aspect of subjective awareness, a conclusion tending to support the strong anthropic principle, although for reasons somewhat different from those usually put forward in support of it. And with the priority of subjective awareness, there is no question of subjective awareness *per se* evolving from a material system since material externality is itself an aspect of subjective awareness.

The major disadvantage of such an approach is solipsism, or rather not so much solipsism itself as solipsistic behavior. If indeed there is a world of other subjectively aware beings as external realities with whom the subjectively aware philosopher must interact as if they have individual external integrity, then any behavior based upon a solipsistic belief will appear psychotic and most likely result in our philosopher's admission to a mental hospital.

So finally we must ask, Is there any solution to the nasty dilemmas that occur both when we assume the priority of external material reality and when we assume the priority of subjective awareness?

V. AN INTEGRATED APPROACH TO THE PROBLEM OF SUBJECTIVE AWARENESS AND MATERIAL REALITY

One way to try to get a handle on this knotty problem is to consider the phenomenon of Absolute Unitary Being, which we briefly

introduced earlier in this paper. As we have described, AUB is a state, usually achieved through intense meditation, in which there is pure awareness with the perception of no discrete reality, the sense of no passage of time, the sense of no extension of space, and without the self-other dichotomy. The existence of AUB is amply attested to cross-culturally in the mystical literature of all the world's great religions as well as by living mystics, whether of the religious or of the secular variety. Furthermore, our physical model of how AUB is generated, i.e., by deafferentation of areas of the parietal lobe, seems to be confirmed in our brain-imaging studies of mature contemplation in Tibetan Buddhist meditators. In short, there can be little doubt that AUB exists, even if it is a relatively rare state.

From the point of view of our concerns here, AUB has an interesting property. Neither during the experiencing of AUB nor upon subsequent recollection is this state ever perceived as subjective. Although it is attained by going deeply within the subject, once it is attained, it is perceived as neither subjective nor objective. Indeed, from a phenomenological perspective, AUB seems to be anterior to either subject or object. Of course, awareness *of something* is clearly perceived to be a subjective state. But as difficult as this may be to understand, the pure awareness experienced in AUB seems to be neither subject nor object when analyzed by the meditator after the fact. Pure Awareness is a philosophical concept that is defined as awareness without content. Pure Awareness is experienced when one is in the state of AUB, but AUB is a specific phase of consciousness. AUB seems to be the only state to which humans have access that eludes the categories of subjectivity and objectivity. If we approach AUB from the stance of giving material reality ontological priority, it can be said to be generated by the simple or absolute functioning of the deafferented posterior superior parietal lobules. If we approach the pure awareness experienced during AUB from the position of giving subjective awareness ontological priority, then we must conclude that Pure Awareness represents absolute reality, in itself neither subjective nor objective, but from which both subjects and objects are derived. Such an approach requires that both individual subjective awareness and external material reality derive from Pure Awareness, as counterintuitive as that may seem.

Thus we must clearly differentiate subjective awareness from pure awareness. In fact, even from the perspective of the priority of external material reality, it seems that the self or conscious ego (as the locus of subjective awareness) has no a priori status but is a practical construct arising from physical evolution. It is beyond the scope of this paper to describe in detail how the Self or the conscious ego may evolve. Suffice it to say that certain brain structures must have evolved before a con-

scious Self could be constructed. Most important among these structures is the inferior parietal lobule, which Luria and others have demonstrated is intimately associated with abstract classification and, more generally, with reification. Thus, as we have noted above, the mind/brain observes its functioning, and this functioning serves as input into the brain. The diverse elements of this mind/brain input are reified by the inferior parietal lobule interacting with the sensory association areas in the same manner that these structures can classify a dogwood tree, a birch tree, and a giant redwood all in the same category which eventually receives the name of *tree* when processed through the language centers. This reification of the perceived diverse functions of the mind/brain is the conscious Self. Once this system of reification is set up, it can operate with infinite redundancy resulting in the Self being aware of the Self being aware of the Self and so forth.

The point of all this is that even from the perspective of the priority of the external material world, the conscious Self is a construct of evolutionary processes, is always aware of "the other" and is unquestionably distinct from the state of AUB. Thus, whether one approaches the problem of the relationship of subjective awareness and material reality giving priority to material reality or to subjective awareness, it is clear that individual consciousness is a secondary reality deriving on the one hand from organic evolution and on the other hand from pure awareness.

To return to the phenomenological analysis of AUB, one must conclude that, unlike individual subjective consciousness, AUB is nonlocal and unlimited. Again, this is counterintuitive from the point of view of Western science. It arises as a necessary conclusion from a phenomenological analysis of the state of AUB and from an epistemological analysis of what the concept of reality can possibly mean in any context. This is as far as we can go with a phenomenological analysis of AUB. However, this ineffable state has tempted many mystics and some philosophers to speculate that the pure awareness experienced in AUB is not only nonlocal but also creative. In such an understanding, not only is pure awareness anterior to subjectivity and objectivity, but it actually creates those categories and possibly the contents of those categories. At this point, one can see the possibility of an externally creating God or ground of being beginning to emerge. Pursuit of this topic, however, would lead us beyond the scope of this paper.

Up to this point, this paper has been a dialogue or counterpoint between the concepts of subjective awareness and material reality. We have seen that subjective awareness can first be understood as deriving from the brain and organic evolution (priority of material reality) and secondly as deriving from a phenomenological analysis of the knowing subject (priority of subjective awareness). We now can see the great

difficulty of relating awareness even to its own primordial machine, i.e., the brain. What, if anything, can we say about consciousness or subjective awareness and artificial intelligence?

VI. CONCLUSION: CAN ARTIFICIAL INTELLIGENCE DEVELOP CONSCIOUSNESS?

If we focus on the phenomenological approach, and especially considering the nature of pure awareness, then advanced computers, like persons, trees, rocks, and atoms, may ultimately be creations of pure awareness. If pure awareness can endow humans, and perhaps animals, with subjective awareness through some algorithm of circuit complexity, then certainly conscious computers are possible. This is not to say that computers must necessarily gain subjective awareness when the requisite degree of complexity is attained, since we do not have the slightest idea how consciousness is generated in the brain itself except for its probable correlation with neural complexity. Of course, anything is possible. If pure, nonlocal awareness creates reality, it does so in a mode and according to criteria that are inherently beyond our ability to know. So from the point of view of a phenomenological analysis of awareness, it is certainly possible for artificial intelligence to become conscious or subjectively aware, although this conclusion is by no means necessary.

If, on the other hand, we approach the problem according to the underlying assumptions of the scientific method or, as most philosophers until the seventeenth century approached it, granting ontological priority to material reality, we are faced with a basic discontinuity between matter and subjective awareness. Using this more familiar epistemological approach, we still do not have the slightest idea of how subjective awareness comes into being, although we may understand the physical mechanisms underlying changes in subjective awareness. With this approach, individual awareness is an evolved entity, but strictly epiphenomenal. We do not have pure awareness as a creative or generative entity underlying everything. This makes our ignorance of how subjective awareness comes into being even more devastating than with the phenomenological approach. Thus, with the scientific approach of naive realism, one must be even more tentative in granting the possibility that artificial intelligence could evolve subjective awareness. Of course one can simply assert, even in the model of naive realism, that when a certain degree of complexity has been attained, subjective awareness will simply come into being—in both artificial and natural systems. This may be true, but the assertion involves a staggering act of faith, more sweeping and more profound than that required by the phenomenological approach to reality. It is still possible that a system of

artificial intelligence may attain subjective awareness, but the gnawing question of why subjective awareness should exist at all profoundly haunts our speculations and tends to undermine scientific faith.

Thus, the same problems that plague our understanding of the relationship of consciousness to the brain also interfere with our understanding of the possible relationship of consciousness to artificial intelligence. As with the brain, so with artificial intelligence. Once a certain degree of complexity is reached in the hardware, subjective awareness or consciousness may come into being for the machine as well as for the brain. But for artificial intelligence as for the brain, there is a fundamental discontinuity between the machine and consciousness. And so awareness may be unknowable in itself, either in its pure form or in its individuated subjective form (i.e., ego consciousness), although we may someday come to know everything there is to know about the neurophysiology underlying the content and phases of consciousness. However much we study the complexities of neuroepistemology, the relationship of consciousness or subjective awareness to the machine, any machine, is a mystery and will likely remain so.

REFERENCES

- Brentano, F. 1973. *Psychology from an Empirical Standpoint*. London: Routledge and Kegan Paul.
- d'Aquili, E. G., and A. B. Newberg. 1993a. "Liminality, Trance, and Unitary States in Ritual and Meditation." *Studia Liturgica* 23(1):2-34.
- _____. 1993b. "Religious and Mystical States: A Neuropsychological Model." *Zygon: Journal of Religion and Science* 28 (June): 177-200.
- Descartes, R. 1911. *The Philosophical Works of Descartes*, vol.2, trans. E. S. Haldane and G. R. T. Ross. Cambridge: Cambridge Univ. Press.
- Gödel, Kurt. 1990. *Collected Works*, vol. 2 (Publications 1938-1974), ed. S. Feferman et al. New York: Oxford Univ. Press.
- Husserl, E. 1931. *Ideas: General Introduction to a Pure Phenomenology*, trans. W. R. Boyce Gibson. London: Allen and Unwin.
- _____. 1970. *The Crisis of European Sciences and Transcendental Phenomenology*, trans. D. Carr. Evanston, Ill.: Northwestern Univ. Press.
- Joseph, R. 1990. *Neuropsychology, Neuropsychiatry, and Behavioral Neurology*. New York: Plenum.
- Llinas, R. 1988. "The Intrinsic Electrophysiological Properties of Mammalian Neurons: Insights into Central Nervous System Function." *Science* 242:1654-64.
- Lynch, J. C. 1980. "The Functional Organization of the Posterior Parietal Cortex." *Behavioral Brain Sciences* 3:485-99.
- Merleau-Ponty, M. 1962. *Phenomenology of Perception*, trans. C. Smith. London: Routledge and Kegan Paul.
- Mountcastle, V. B. 1976. "The World around Us: Neural Command Functions for Selective Attention." *Neuroscience Research Progress Bulletin* 14:1-47.
- Mountcastle, V. B., B. C. Motter, and R. A. Anderson. 1980. "Some Further Observations on the Functional Properties of Neurons in the Parietal Lobe of the Waking Monkey." *Behavioral Brain Sciences* 3:520-29.
- Newberg, A.; A. Alavi; M. Baime; E. d'Aquili; D. Duncan; and P. Veloso. 1995. "Cerebral Blood Flow During Intense Tibetan Buddhist Meditation Measured by HMPAO-SPECT." Annual Meeting, American Neuropsychiatric Association, Pittsburgh. (abstract)

- Newberg, A.; M. Baime; E. d'Aquili; D. Duncan; and A. Alavi. 1995. "HMPAO-SPECT Imaging During Intense Tibetan Buddhist Meditation." Society of Biological Psychiatry Annual Meeting. Miami, Fla. (abstract)
- Penrose, R. 1994. *Shadows of the Mind*. New York: Oxford Univ. Press.
- Wilber, K. 1975. "Psychologia Perennis: The Spectrum of Consciousness." *Journal of Transpersonal Psychology* 2:105-32.