

[[My first paper on consciousness. Written in 1989 when I was a 22 yr. old MS student studying Artificial Intelligence. This was the term paper for my first philosophy class. The specific project proposed here was never developed further, and looks very fanciful now, but it is interesting to see how close to the surface lie the "big" ideas that eventually became *A Place For Consciousness*.]]

Consciousness As A Physical Property And Its Implications For A Science Of Mind

**Gregg H. Rosenberg
Artificial Intelligence Programs
111 Graduate Studies Research Center
The University of Georgia
Athens, GA, 30602**

Abstract

As the view that the mind has a physical cause becomes increasingly more difficult to refute, both philosophy and science must face the fact that having experiences, qualia, consciousness in short, is simply not deducible from within our physical theories. Indeed, all the power physics shows for qualitative explanation is adduced from outside the actual formality of its theories. Our physical theories describe vibrations and stochastic correlates of motion, and there is no principled way to explain awareness or the existence of experiencers by mere vibrations or motions. The problem arises because the objectivity of the language of physical theory is antithetical to the subjectivity of consciousness. The gap between them can be understood analogously to the gap between "is" and "ought" reasoning in ethics. One solution may be to bypass formal languages that attempt to purely deduce consciousness from without, and instead explain it using a pseudo-poetic language that can withstand both physical and introspective interpretation. This paper introduces such a language, and it uses the new language to define an "Ontological Principal."

Preface

This essay is an attempt to fit consciousness into a physical worldview by expanding our ideas of the nature of the physical world to encompass more than just the descriptions of physics. This is not a reductionist argument in the sense put forth by Fodor in *The Language of Thought*. Such arguments from the special sciences to physics are of the form $1 \Leftrightarrow 2$, where the left side of the bi-conditional contains the laws of the special science and the right side contain some kind of bridge laws that lead towards the laws of physics. Fodor gives a convincing argument as to why we should not expect such a reduction for cognitive psychology. The strategy taken here is to explain consciousness by immersing physics inside a larger and less formal view of nature able to withstand a dual interpretation from both introspective and objective points of view.

The topic of this paper is consciousness, the qualitative nature of experience, which is not often addressed fully by modern cognitive psychology. Cognitive psychology's relatively shallow treatment of the issue of consciousness is not surprising because many workers in the discipline accept a strong view of functionalism and take the computational metaphor for understanding the mind to be more than a metaphor. As Fodor says in his introduction to *Representations* (1981)

" And while it is arguable that what makes a belief or other propositional attitude the belief that it is is the pattern of (e.g. inferential) relations that it enters into, many philosophers (I am among them) find it hard to believe that it is *relational* properties that make a sensation a pain rather than an itch, or an after-image a green after-image rather than a red one...It seems to me, for what it's worth, that functionalism does not provide an adequate account of qualitative content; that, in fact, we have no adequate account of qualitative content."

Thus, while this paper wishes to begin us along the road to an adequate account of qualitative consciousness, it begins by suggesting that we subsume the laws of physics into a less formal, but empirically constrained, understanding of the physical world from which we may, indeed, reasonably speak about physical objects having sensations, experiences and so forth. Particularly, this essay may be seen as an attempt to elucidate dual aspect theory with some emphasis on the problem of how simple physical objects can combine to make complex physical objects that have unity of experience. Nagel (1986) states that the combination problem, along with panpsychism, are the two great problems facing a dual aspect theory. I believe making inroads on the combination problem will

simultaneously alleviate reservations about panpsychism.

There are two key ideas to this elucidation, neither put forth dogmatically but only as tentative pointers down a road worth journeying. First is the idea that such an understanding of the world will be in a dual interpretation theory. A dual interpretation theory is single narrative structure that can be interpreted as referring both to the objectively observable characteristics of matter and the subjective experiences of that same matter. The term used in the essay is "pseudo-poetic" because of its relation to literary devices such as allegory, metaphor, simile, etc.

The second idea is that a bridge theory cannot be built from consciousness to physics because there are no good candidates in our physical lexicon to link consciousness to. Theorists in a qualitative philosophy/psychology are not standing on one edge of a canyon but, rather, are looking down the precipice of a ravine. All other successful reductions have had properties of matter to bridge to. The reduction of the temperature of a gas required kinetic energy, the pre-eminence of phlogiston was undermined when flame could be linked to the energy in chemical bonds, etc. The absence of a concept to bridge from in physics is especially evident when physicists themselves write about consciousness. Roger Penrose (1989) attributes to consciousness otherworldly powers for truth divination, and a connection to a Platonic realm. John Von Neumann and Erwin Schrödinger call upon consciousness to create reality by collapsing the wave function in quantum mechanics. David Bohm (1984) suggests that consciousness is a hidden part of the implicate order somehow made explicate by the mind. I feel safe in concluding that if a candidate for a bridge theory from consciousness to physics existed, one of these theorists would have spotted it.

However, the absence of a notion for deriving consciousness from physics is, of course, not the same as the absence of properties of the physical world from which consciousness may be explained. We needn't draw an identity between physics and physicality. The response in this paper is to try to expand our notion of physicality as, for instance, Bohm does when he postulates the Implicate Order. Independently of accepting Bohm's physics, we may note several experimental and theoretical results that have prompted Bohm and other physicists into their alternate interpretations of the physical world, and we may agree that they are usefully like the properties matter would have to exhibit for us to sensibly maintain that matter may be responsible for consciousness.

First, the independence of what seem to be separate objects has turned out to be a relative matter, with objects appearing to be more and less independent of one another in different contexts. The degree of independence runs the gambit from the point where some objects exhibit a

baffling unity of behavior even when vastly separated in space to the point where objects seem completely oblivious to one another. Moreover, it seems easier and more sensible to derive relative independence from fundamental unity than to derive a relative unity from a fundamental independence. This is important, for it means that we needn't derive the unity of consciousness from the actual separateness of the physiological parts but may assume consciousness is evidence for an ontological unity in the parts from which we then must explain why they appear separate to the senses ? an easier task.

Second, Ilya Prigogine (1984) has produced results that suggest that irreversible, time dependent, processes exist in nature as a fundamental attribute of the physical world and not merely as the result of our own ignorance, or as an illusion. Additionally, systems exhibiting irreversibility arise when matter is in a state far from equilibrium, such as with organic beings.

Third, large scale systems, called quasi-crystals, have been discovered which are suspected of organizing non-locally.

Fourth, all three of the previous discoveries rely on non-local descriptions of matter, descriptions in which, in Bohm's words, "One discovers, instead, both from consideration of the meaning of the mathematical equations and from the results of the actual experiments, that the various particles have to be taken literally as projections of a higher-dimensional reality which cannot be accounted for in terms of any force of interaction between them." He then goes on to point out that such results make sense only if we think of the apparently separate phenomenon as being different *aspects* of a higher dimensional unity.

Why are these results of interest to those worrying about consciousness as a physical property? It is because they suggest that some of the introspective experiences we have of ourselves that have been alien to our physical understanding of the world may indeed have an echo in the observed behavior of physical systems. In the first case, the experience of unity of consciousness has an echo in the unified behavior of separated particles. In the second, the experience of time flowing unidirectionally - an experience starkly contradicted by physics for hundreds of years - has an echo in processes of complexity and growth. In the third case, we see that these behaviors are not necessarily limited to the quantum realm but may be echoed in large scale phenomena. In the fourth case, we see a betrayal of a kind of unity that is more fundamental than extension in space or time and can manifest itself in phenomena extended in space and time. We learn to distrust the individuation criteria of our senses.

I do not wish to make too much (or too little) of these behaviors and the importance they have for a theory of consciousness. I do not think it is wise or justifiable to argue for meaningful quantum effects in brains, or to

generalize from an isolated example like quasi-crystals to other phenomena too readily. A theory of consciousness should not tie itself to any speculative physics, but still it should pay attention to the phenomena that are causing physicists problems. These examples are important because they show that the physical world can surprise us, and surprise us in ways that remind us of ourselves. Although the theories which describe these processes, as they stand, may not provide a bridge to consciousness they nevertheless may have echoes in brains and nervous systems, echoes which in turn are echoed in our experience of ourselves. They may betray an insideness to matter that somehow becomes the insideness we know as feeling, awareness and consciousness.

Finally, I should say that the ideas in this paper are more compatible with a process metaphysics than a predicate metaphysics in which the world exists fundamentally as objects with properties. One of the hurdles a dual aspect theory of consciousness must overcome is the fact that what appear to be many somehow combine to become one. I feel consciousness and the existence of objects are best reconciled if objects are seen as contingent and transient results of a continuing movement. Thus, Whitehead's conception of organisms as the nexus of occasions allows us more readily to combine a many into a one than any conception of the many as objectified does. Just as waves may merge and become irretrievably entangled but particles may not, in a process unity may be achieved from many processes as a true merging into another existence, but I do not think an object metaphysics can achieve this same feat. Regardless, this essay describes an object-centered world because it is an attempt to unite consciousness with the physical world as we know it. Even if on a meta-physical level the world is process, on a physical level it is perceived and understood as a collection of objects, and to write about it any other way would be obtuse.

"Scientists would be wrong to ignore the fact that theoretical construction is not the only approach to the phenomena of life; another way, that of understanding from within (interpretation), is open to us"

- Herman Weyl

I know something not many will be surprised to hear, yet I will bother to say it anyway. The evidence mounting in the neurosciences strongly suggests that for every mental event there is indeed a corresponding brain event, and brains are governed by the laws of physics. Taken together, these two conclusions strongly imply that minds are physical phenomena. Although dualism is not logically incompatible with the facts, upholding it increasingly seems curmudgeonly, and there are no longer many philosophical camps willing to carry the banner. The new difficulty is now, of course, explaining how mental phenomena can be compatible with our physical theories. There exist empirical programs attempting to map physical and mental phenomena by thinking of the brain as an information processor. The thesis of this paper is that such efforts, although fruitful, can only meet with partial and aspectual success. The problem is not one of sheer complexity, although this is surely a prodigious problem. The root of the problem is profound, and reaches back into our most primitive notions of naming. There is a basic incompatibility between the world implied by the language of our sciences and the existence of consciousness. I will try to defend the notion that a science of mind must explain consciousness in physical terms, and can only do so by inventing a new theoretical language.

The purpose of this paper is to suggest that we can invent a language for re-interpreting the results of the various fields into a more inclusive and universal framework. The new effort I would like to see would not obviate the need for research into the structural aspects of brain and mind, but would incorporate them into a more inclusive framework, re-interpret the results to give an inside-out understanding of ourselves. It would be both a scientific and literary (interpretive) enterprise and it would address the tension between P1 and C1 below:

P1 Minds are natural phenomena.

P2 Minds are simply instantiations of algorithms.

P3 No natural property is essential to the instantiation of an algorithm.

C1 No natural property is essential to being a mind. (P2 and P3)

The proposition that something may be a natural phenomena and yet have

no essential natural properties is precariously tense. Yet, we may go further:

P4 Consciousness is a natural property of minds.

C2 Consciousness is not an essential property of minds. (C1 and P4)

P1 and P3 being untouchable, it seems we must deny either P2 or P4 to avoid C2. In neither case do we explain just what, then, it *is* that consciousness is. In redressing this tension, I will attempt to assign a place to consciousness in our own minds by seeing the fundamental elements of it in the world about us. In attempting to identify the principle behind consciousness, we will speculate that digital computers cannot be conscious (denial of P2, the thesis of strong AI). In speculating about the new language a science of mind must invent, we will find that aesthetic criteria and truth criteria will both play a role. In short, a science of mind must bring together fractured perspectives.

"Do you believe that absolutely everything can be expressed scientifically?"

- **Hedwig Born to Albert Einstein**

"Yes, it would be possible, but it would make no sense. It would be description without meaning - as if you described a Beethoven symphony as a variation of wave pressure."

- **Einstein's reply**

No alarm should be raised at the mention of inventing a new language in which to theorize. Science has done it before. We call it mathematics. Because the terms of everyday life are ripe with consciousness and experience, analogy and vagary wholly unsuitable for the kind of analysis science wishes to undertake, we dispensed with them. In their place we substituted the beauty of $E = mc^2$, which does not need to state whether things are matter or energy but is able to supersede the disjunction experience would infer, simply stating that they are elegantly equivalent.

We cannot continue with a strictly formal language for theorizing in a science of mind because, as Anthony Nemetz argues, we first began using these languages as a way of minimizing consciousness in our descriptions. We did this for good reasons. Consciousness was immersed in time, but our theories needed to be general, to transcend the flux. Therefore we required a language removed from time, whose terms could be related in an eternal sense. Mathematics served this purpose well because its emphasis is on the connectives between terms, on the unchanging structure and not the continually shifting particularities. Mathematics protects our insights by

freezing them in a stillness of rule so that we may inspect them at our leisure; a stark necessity caused by the questions we were asking growing too big for the language of experience that we were using. As Nemetz states, "Because mathematics has no concern with external references or existential conditions, mathematical concepts become in effect, concepts of concepts or models of models of the world. This means that mathematical laws, especially those of Boolean algebra may claim to function as laws of thought." Yet consciousness is not existentially indifferent. Consciousness is, rather, the very expression of existence. The call for a new language that I am now making is no more than a simple plea of recognition, recognition that the questions we are asking have once again outgrown the language that we are using, mathematical logic.

How do we know that our questions have outgrown this language? The answer lies in noting that consciousness and mathematics may simply express two entirely different, but complementary, aspects of the world. Mathematics expresses the relations between terms, but is formally indifferent to the existence of the terms. Quite oppositely, consciousness is the mind's device for expressing existence directly. That is, consciousness is that thing which formal systems must take for granted; it is the "interpreted" part of "interpreted formal system" with the result that consciousness cannot be deduced from within but must always be supplied from without. Consciousness must decide the truth of the formal system and thus cannot be proved within it. Therefore, formal systems are imprisoned within a dogged aspectualness. Our mathematics may describe our physical world perfectly adequately without any mention of qualia. And our logic's of qualia need not incorporate physicality.

A canonical example makes this point well. Philosophers of science have stated that the color red has been explained by the theory of electromagnetic radiation. But nowhere in the theory can one deduce the qualia of red. We know that this theory explains why some things look red only because we have experiences of red in our lives and can willfully bridge the gap between what the theory describes and what we consciously experience. An evolutionary descendant of humankind, who can no longer experience the color red, will gain no insight into that conscious experience by our theory of electromagnetic radiation. Redness is not explained by the theory, it is adduced into it. A science of mind is thus faced with the choice of continuing to use the traditional languages of the sciences, and thus assume one of the phenomena it is supposed eventually to explain, or to invent a language for theorizing that includes terms of consciousness and their attendant imprecision.

Clearly, merely from the use of language one cannot get a theory to induce experiences of consciousness that are not indigenous to the knower.

The goal must be more modest. We may begin by noting that formal systems are "about" while consciousness "Is." Our inability to deductively bridge the gap between "about" and "Is" leads to the *ad hoc* nature of our electromagnetic theory of red. The closest approximation we have to a language that expresses "Is" is our introspective language, the language we use to describe our immediate experience. Its most effective use in conveying consciousness is in literature, especially metaphor. Yet this language is far too vague and idiosyncratic to be of much use in serious theorizing.

The real problem in trying to explain consciousness physically is that all our ideas of the physical world are arrived at by reasoning about things that we are outside of, while all our ideas about consciousness are arrived at by reasoning about ourselves from the inside. This inside/outside distinction naturally suggests a dual aspect theory. For instance, if Joe is facing Fred's eye Joe see's a football shaped white thing with a colored circle moving around in it, but in the same situation Fred's eye sees Joe's face. It is just a different experience to *be* Fred's eye than it is to *look* at Fred's eye. How can such a situation be spoken about? We may steal a compromise from the arts. Whereas mathematics is existentially indifferent, Nemetz notes that the languages of the arts are existentially ambivalent. They simultaneously may refer to many things. We need a language that can refer to both aspects of Fred's eye simultaneously. Our solution may be to invent a scientific language that can be interpreted to refer both to physical relationships and introspective experience in a meaningful way. That is, a theory of consciousness needs to be stated in a pseudo-poetic fashion. By "pseudo-poetic" I mean a single narrative structure that can be interpreted in two ways despite apparently unrelated subject matters. As with metaphor, we can use this as a linguistic device for portraying an underlying unity between the apparently diverse subjects. By scientific, I mean the narrative purports to explain natural phenomena, is objectively understandable, is an attempt to interpret empirical knowledge, and is amenable to experience. As a note, to those who would protest that a theory that is not formal is not *really* scientific I would provide natural selection as a counter-example. It is rigorous, empirical and explanatory without being formal.

Finally, I need to tell you what exactly I am including under the term "consciousness." This term, I am using in a different sense with which you might not be familiar. Often, the literature uses consciousness to mean merely self-consciousness (as Minsky, 1986, seems to), but those are two terms I would like to differentiate. Also, theoretical frameworks usually invoke some concept of memory organization or access to memory (e.g. Tulving, 1985) and I will ignore this work as describing a special property of a more general phenomenon. Consciousness is primarily the qualia involved

in existing. In humans, consciousness is having an image or a sound or a taste. Self-consciousness is a complex of consciousnesses interacting in a certain way. I will discuss both these ideas shortly.

"It is a good deal easier for most people to state an abstract idea than to describe and thus re-create some object that they actually see ... But there's a certain grain of stupidity that the writer of fiction can hardly do without, and this is the quality of having to stare, of not getting the point at once"
- Flannery O'Connor

I maintain that a good theorist must be a good writer of fiction, and so we will begin to look for consciousness with a certain grain of stupidity. Certainly, not everything has consciousness in the way I define it here. Rocks do not see pictures, the sea does not taste its own salt, and trees are not tickled by the wind. Yet they are made of the same stuff as us, governed by the same laws. Somewhere in there must be, if not consciousness, then the foundation of consciousness, the hint of it. To find it I am going to have to answer a thick-headed question.

There is a bar near where I live that has pool tables at the back. I am not a very good player, but I sometimes go there to relax and shoot a few games. One day, I was struck with surprise by the fact that when one ball hit another they careened off in response. I honestly wondered why. The "why" I wondered did not concern angular momentum or mass. It was a pre-nascent, fundamental, how-the-heck-do-they-know-they've-just-hit-into-each-other kind of "why." In information processing terms, signals had been sent between the atoms of the two balls. The signals contained information about angular momentum, mass, and so forth, and each ball responded appropriately according to the fixed physics of the matter. This kind of explanation did not answer my question at all, though. I was astounded by the fact that the signals could even be sent and received, that atoms could communicate these forces. What a strange world it is where two things can actually affect each other!

I began to try thought experiments. How would a universe be if things could not affect each other? A few scenarios popped into my head. First, I thought of utter chaos, but even in chaos things react to one another. Then I imagined ghostly images passing through one another unaffected, but this involved overwhelming space/time problems such as two objects occupying the same space/time. Finally, it simply seemed that such a universe could not *be* at all. Each thing would be complete unto itself, a reality utterly still,

floating, full of darkness, dead. Fundamentally, in order for there to even be a universe, of any type, the things composing it must be *receptive* to each other in a very simple sense. This implies that in everything there is an appropriate sense of being that thing ? namely, the sense in that it is connected to other things and can receive knowledge about them. I call this quality receptivity, and I believe it is the fundamental manifestation of consciousness in physical phenomena.

"We feel acutely our separation from that primitive perception of the world. Unlike our animistic forebears, we are not at home in the world ... And we diagnose the case to favor our own symptoms: they were the victims of a delusion that projected consciousness upon a dumb, drooling, cretin of a cosmos"

- Virginia Stem Owens

In the previous paragraph I used the term "knowledge" in a very controversial way, in reference to atoms. The heckles on the tongues of quite a few hard-nosed philosophers and scientists should have been raised. I do not wish this to be a strange notion, merely an inside-out notion. From the outside, this thing I called "knowledge" is usually referred to as "information" or "phase state." However, one cannot get at consciousness purely from the outside, and attempts to describe behavior in these traditional ways inevitably leave the subjects, in some sense, dead. Yet, I am not content to leave this renaming as merely literary anthropomorphism. If we are going to be able to work with terms, we must define them in a way that allows two things. One, our minds must be able to work with them in a facile way, and, two, all minds must get reasonably similar meanings from the definitions. Formal systems meet the second criteria exceptionally well because the terms are interpreted very literally and are set up in precise relations to each other. This is one reason they have been successful as the language type of science. Additionally, for certain kinds of thinkers, they also do an exemplary job of fulfilling the first criteria, hence the phenomenal success of the sciences.

Unfortunately, I have already taken the position that a formalism can only lead to an *ad hoc* explanation of consciousness. I will still venture to define the terms so far discussed:

Thing - The property some things possess of being ontologically individual.

Receptivity - The property in Things, living and non-living, which leaves them vulnerable to being changed by things.

Knowledge - The entire potential in a Thing that gives it the ability to change other things.

Introspectively, Thing is meant to capture the unity of experience, Receptivity the possibility of feeling and Knowledge the reality of presence. Physically, the approximate physical description of Knowledge and Receptivity, together, comprise the canon of physical laws, and Thing encompasses all classes of entities in physics. The purpose of defining these three terms is to ground consciousness in the quality which allows participation in the universe and manifests itself as the quantitative properties physics describes. This participatory quality must emanate from and resonate in the insideness of Things, and it logically follows that some Thing that participates must have an outside effect. This effect is what is the outside aspect of the Thing. Also, I stated earlier that consciousness is an expression of existence. The definitions of Thing, Receptivity and Knowledge can be seen as the conditions of existence. That is, in order to for us to say properly that an entity exists it must be individual, effective and responsive. Therefore, both inside and outside aspects of entities are referred to. Finally, Receptivity and Knowledge together can be seen as portraying the underlying connectedness of Things. Our common sense ideas of cause and effect between independent objects are an approximate theory of this connectedness, a theory adequate only in a limited range of cases.

Being - The manifestation of Knowledge. The approximate physical description of the Being of a Thing at any given moment is the result obtained by measurement. The introspective interpretation is sensation isolated in time. Being is the description of a Thing grounded in time, of the Now that troubled Einstein so, and it is also the particular aspect of a Thing's Knowledge that is currently available to be received by other Things. The more of a Thing's Being that is available to another Thing, the more connected those Things are. In this way we try to portray the relative nature of the independence between existing objects. We can imagine that in some contexts that a society of Things may become so entangled in one another's Beings that they unify. In other contexts, they are so disentangled they appear wholly independent.

I wish also to define:

Meaning - Knowledge in transmission through Being. Meaning is meant to be considered as a physical phenomena involved in the creation of consciousness. It has no specific counterpart in the physical sciences, but is always assumed. As a society of Things becomes very entangled, we shall say that their Meaning becomes very rich, while a set of Things that are not entangled generate poor Meaning. When a society of Things are unified, then they generate a single Meaning in which their single experience is felt.

Receptive System - A system whose parts cooperate as the result of their collective Meaning. All conceivable systems are receptive on some level.

Contingent System - A system whose parts are coordinated as the result of an incidental pattern they trace in space and/or time, and do not require any specific Receptivity, Knowledge or Meaning. Contingent systems are systems which appear to the senses to be unified but, in fact, have a very poor Meaning. For example, automobiles, and many other artifacts, appear individual but are not ontologically unified, only contingently coordinated.

Although metaphysical sounding, the terms proposed are motivated by a desire to explain what seems to be empirically true about both consciousness and objectively observed physical systems, and thus are meant to refer to physical phenomena. They each have a dual interpretation, and were chosen because I felt that they were suggestive of their common sense counterparts (although not identical with them). Their counterparts are most often treated as abstract entities and have been the result of much discussion during which they have acquired many meanings. But the terms of this language are not meant to refer to abstractions and are going to be used differently here than their counterparts are in the philosophic and scientific literature. As a way to distinguish, I shall capitalize the terms of this language when I use them. I have already made use of this method of reference in the definitions themselves. Thus Thing, Knowledge, Meaning and Receptivity are not the same as thing, knowledge, meaning and receptivity. Pointing out the difference between "Thing" and "thing" should be illuminating here.

The senses identify and experience individuals as degrees of coordination among parts and thus create sensory individuals which may not be ontological individuals. Coordination, here, is a continuum, and the exact placement necessary to classify something as an individual is fuzzy, and exists as a relative concept, a basis of comparison. We say that A is not B

because the intra-coordination among the parts that allows us to recognize A and B is far greater than the inter-coordination between the parts we ascribe to A and the parts we ascribe to B. Notice however, that in the definition of Receptive System the word cooperate is used to describe the actions between the parts, but in Contingent Systems the word coordinate is used. Cooperate is meant to connote an intrinsic tendency in a Thing to relate Meaningfully to another Thing in the given way. Coordination, on the other hand, can result merely from happenstance, because things happen to be located together or move together in a certain way.

Since our senses receive abstract information about Things, pattern information, and not information about the reason the patterns exist, a coordinated system and a cooperating system are individuated just the same by the senses. That is, the senses individuate without regard to ontological status and it is to this sensory individuation that we normally apply the word "thing." For example, in ordinary thought we think of paintings as "things." However, in the empathetic language proposed here, paintings are not ontologically basic "Things" because the fact that many millions of pigments of paint are lying next to each other on a surface is simply a Contingent and not a Receptive fact. Our senses and not its basic ontology individuate a painting. I shall later apply a similar argument to the case of computer consciousness, but we shall see that the answer is less clear in that arena.

These terms serve chiefly to give form to an Ontological Principle. The **Ontological Principle¹ being proposed is that there are certain Things in nature which have a basic, if transient, existence as ontological individuals, as opposed to sensory individuals, and that these Things can sometimes come together to entail the creation of another Thing. For the sake of concreteness only, we can imagine the basic Things as being the elementary particles of physics. These Things have the quality of Receptivity. In our empathetic thinking we can ascribe to them a certain awareness and connection to other Things which allows them to be affected by other Things and to react accordingly. The affecting and subsequent reacting are acts of Meaning. The Ontological Principle states that sometimes Meaning not only affects the Things involved, but also leads to the creation of an altogether new and single ontological Thing. This new Thing is not in place of the old Things, but in addition to them, and has its own Receptivity, Meaning, Knowledge and Being. The Things which**

¹ Not to be confuse with Alfred North Whitehead's ontological principle. Whitehead's principle is an attempt to incorporate both final and efficient causation in his meta-physics. The ontological principle put forth here is meant to portray a natural principle by which a multiplicity may also be a unity.

make it up form a richly Receptive system.

To be more concrete, let us consider a Receptive System like the chemical reactions in the brain. It consists of parts reacting according to their unique and intrinsic causal properties. The particular character of an interaction is determined by the Being of each Thing involved, the Being of these things is varying significantly through time, and thus their interactions are irreversible in time. In our empathetic language we shall term these interactions to be, for each, expressions of their Knowledge through Being. This is Meaning. Sometimes parts become entangled in a symbiotic relationship in which there is not only Meaning, but also that Meaning is directly responsible for the maintenance of the relationship. The relationship itself then becomes both an expression and an origin of the Meaning of the system ? a single process with an emergent Meaning. When physical systems are in this symbiosis, the natural result is for them to express this new emergent Meaning as a single Meaning. This is the Ontological Principle at work. Thus a new Thing is created. If these symbioses become complex enough, multiplied again and again, the expression of Being for this Thing also becomes more complex and takes the form of images, tastes, sounds, pain, felt emotion ? human consciousness. At other times the interactions between Things are cursory and poor, they do not become properly entangled and a new Thing is not created. Since only Things may be conscious, systems which consist of only these "poor" interactions would not be Things and not be conscious. Thus we may say that a brain, which has rich Meaning, is conscious, but that a painting has poor Meaning and is not conscious.

Naturally, we would expect the consciousness of each Thing to be as unique and varying as the Meaning involved. We would also expect it to be of a universal "character" as the Being involved is an instantiation of Knowledge, and the underlying Knowledge responsible for different Beings may be similar or even identical. In relationships of small complexity, such as molecular structures, the expression of Being must be unimaginably simple and it is probably best to continue to refer to them as simply Receptive and not conscious. In very simple things, like quanta, the variation in Meaning may be so insubstantial that it makes sense to speak of different Things as being identical. But in all symbiotic relationships of considerable complexity there is some sense of qualia, for qualia is simply an expression of Being some Thing. We may quip that consciousness is the Knowledge Representation of organic beings.

This explains why it is immersed in time, for it is expressing the continually shifting Meaning of its parts. It is an expression of continually re-emerging Being.

So the Ontological Principle is a two part invention. In the first part, different Things must communicate their Being to one another in a way that the Meaning preserves the Being of the communicating Things particularly well, allowing their Beings to become entangled with one another. Since Being is historically determined and constantly changing then Being, once faded, cannot easily be reproduced exactly. Necessary evidence of this would be irreversibility in time. That is, the Things may be particularly precocious and resistant to manipulation. In the second part, the communication must result in a symbiotic relationship between the Things. Prima facie evidence may be self-organization and perseverance in time. The result is an ontologically individual Thing. Prima facie evidence of ontological individuality may be non-local behavior in space or time; photons in the half-silvered mirror experiment or quasi-crystals being popular examples. Speculatively, epi-genesis in embryos may be another. Finally, it makes sense to say of any Thing that there is something that it is like to be that Thing, and in Things organized complexly enough to be minds, consciousness is the way of expressing Being. And so consciousness is a (nomologically) necessary property of being a mind.

By relying heavily on ontological considerations, this account, rough as it is, ameliorates some of the more unsettling panpsychic implications of dual aspect theories. For instance, we do not have to ascribe consciousness to rocks or door knobs because we can deny that rocks and door knobs are ontologically individual. In general, we may distrust the individuation of our senses gravely and assert that *most* of what our eyes and ears individuate for us are not actual Things but only things. Indeed, this seems only sensible. In addition, while we cannot avoid ascribing proto-mental properties to odd Things like, perhaps, carbon atoms, we needn't be too upset by it. The character of its Receptivity and Meaning are going to vary greatly with the character and complexity of the Thing that it is. The nature of what it is like to be a carbon atom is inconceivable, probably quite boring, and certainly nothing like what we think of when we think of consciousness. Of necessity, we will simply leave this proto-mental quality acknowledged but unspecified. Filling in more detail would require that one be a carbon atom. Furthermore, this qualitative nature may be a bald fact and we needn't attach moral or mystical importance to it. There need be no panpsychic

"Cult of the Carbon Atom."

Comparing the ideas involved in the workings of the Ontological Principle to those in the workings of an accepted, well understood natural theory that has appeared in discussions of the mind/body problem and philosophy of science may help.

"----->" should be read as "plays a role similar to"

	THEORY OF CONSCIOUSNESS	THEORY OF TEMPERATURE
Physical Entity	Thing ----->	Molecule
Natural Property	Receptivity ----->	Kinetic Energy
Considered	A group of -----> Things	A gas cloud
Behavior	Rich Meaning ---->	Random Collisions
Context	Symbiotic -----> relationship	Defined space
Natural Principle	Ontological ---->	Conservation
Invoked	Principle	of Motion
Result	Thing ----->	Temperature

The new Thing has a more complex, and qualitatively different, Receptivity, Knowledge, and Meaning than the original Things. The analogy between temperature and Thing reduction is suggestive and meant to be a psychological help but is not perfectly symmetrical. One difference between the reduction of temperature and consciousness is that a Thing produced by the Ontological Principle is a phenomenon in nature with properties while temperature is strictly a property. Also, the arrow connecting the entries under "context" and "behavior" might also be interpreted as "contrasted to." They are contrasted in the sense that the rich Meaning involved in creating consciousness precludes the kind of averaging that allows us to treat the collisions in a gas cloud as random. Also, the symbiotic relationship between the parts of a Thing define the space *for* it and us, but the observer can define the space over which to average out a temperature.

Finally, what of self-consciousness? It seems that consciousness is having a qualia and self-consciousness is knowing that you have a qualia. Remember, being a Thing does not entail that your parts are not also Things. A single organism can be made of many Things, and if the organism is complex enough each one of these Things may have the status of consciousness. This Thing which we call "I" is a reflection of lower level consciousnesses, not a re-statement of them, and the reflection is of a very complex organization of already conscious Things. Once this organization

allows for memory, then the new Thing may become self-conscious merely by comparing its consciousness through time and remembering the stability. It is the information processing capabilities of the human mind which allow for self-consciousness. Thus self-consciousness is not a necessary corollary of consciousness, but a special quality of certain types. In fact there is no reason, in principle, these symbioses cannot be occurring on many levels, and each consciousness in us is self-conscious with respect to the Things which make it up. Consciousness is an expression of Being. Self-Consciousness allows the experience of Beings.

**"Consciousness ... In this book, the word is used mainly for the myth that human minds are "self-aware" in the sense of perceiving what happens inside themselves"
? Marvin Minsky**

The strict insistence that any scientific explanation of the mind must be formal has led some to more or less deny the existence of consciousness. Actually, in the above case, it has led to a refusal to recognize what it is and then a denial of something that it is not. Consciousness is a perception that we exist and any relation it bears to " what happens inside" ourselves in an objectively described information processing sense, is accidental. A digital computer, in fact, is a Contingent System, and I shall argue that it cannot be conscious. Therefore, it is not surprising that researchers convinced that computers can be minds would diminish the importance of consciousness. The principle by which a Receptive system can also be termed a Contingent System is very difficult to pin down. However, the ideas presented thus far do not live and die on our ability to find an exact boundary for classifying Contingent Systems. *We know* that some physical objects are conscious, and we strongly believe that *many* other physical objects are not conscious. This alone is reason to believe there is a physical difference of *some* kind in systems. The ideas presented thus far gain or lose their plausibility on the likelihood that they pick out the right kind of differences and not the rigor with which they can measure the differences.

Computers provide an enlightening case for exposing the attractiveness and inadequacies of our understanding of what has thus far been said. The situation we will imagine is simply that mankind has invented a digital computer that shows the same signs of intelligence that human beings do. The question is then, "Is this computer conscious?" Or, put another way, "Obviously, this computer is a sensory individual, but is it also an ontological individual? Is it more like an automobile or painting, or is it more like a brain?" The line I take is exactly Searle's argument against

strong AI. Whereas I do not think Searle's arguments convict the thesis that computers may one day be intelligent (where "intelligent" is interpreted behaviorally), I do think it applies to conscious states. I shall say that digital computers cannot be conscious because the physics which govern their operation do not involve the physical objects in the right kind of relationship. The causation is all wrong. The discussion is an a priori argument all the way, as I believe any speculation of this sort must be. Consciousness is just not empirically observable in anyone but oneself. Of course, I cannot provide conclusive arguments either way, nevertheless, I will try to give suggestive arguments.

Let me emphasize that I am not addressing whether a computer can ever be intelligent. At one time I think intelligence was a word that had connotations of consciousness, but I think it has gradually been appropriated by the sciences and turned into something denoting performance criteria. I think one common definition given in psychology these days is "competence." Minsky more or less decides it is outmoded, and most discussion revolves around some kind of "problem solving." Whether a computer can ever meet the performance criteria that would make us agree it is intelligent is a purely empirical question. Conscious minds do sometimes meet these criteria, and other things may also. I can easily imagine us someday distinguishing between conscious intelligences and non-conscious intelligences.

According to what has been said here, the consciousness of a conscious intelligence is an expression of the Meaning in the system. In a system like a digital computer, the behavior of the parts is the result of the abstract logical design of the system. Each gate seems to act relatively independently and reversibly in time. When a circuit gate opens or closes, it is not the result of a unique communication between gates, but rather because it is part of a certain sequence that has evolved through time. The gates can be gates on a ranch with people walking through them rather than electrons flowing. There is no significant change in the Being of a gate because it is part of a certain program or because a certain person walks through it. The openings and closings it goes through occur just the same if it is part of an intelligent program, a dumb program, or flapping in the wind. As Searle says, programs can be executed by water pipes being turned on and off. This is much like what digitization is.

The rigid digitized nature of today's computers should block consciousness by forcing its parts to communicate mostly circumstance with one another, and only a trivial amount of Being. As prima facie evidence of this, I would like to contrast the fact that a Von Neumann computer can be programmed to execute any program that can be logically specified, while a brain has only one program, the one dictated by physical laws. Additionally,

note that in order to make a human being the process in the womb must be protected, whereas to make a computer the process must be overseen. This is because people happen as the result of the intrinsic nature of the Things comprising them. Computers are an organizational happenstance, and it is not in the nature of the Things comprising them to cooperate. We merely rig them so that they are coordinated.

Perhaps a better way to say this is to note that if we know what state a computer chip is in, then if it is damaged we may melt it down and reconstruct it in the exact same (for our purposes) state. However, now let us assume that the brain is running a computer program and we know the exact program state the brain is in. If the person dies, we cannot melt the body or brain down and remake it into the appropriate state. The process is just not reversible or retrievable. Why is this? One may be tempted to think the answer is trivial, but it is not. The physical components that make up a person are in a radically different kind of relationship than the physical components that make up a computer chip. They are not at all passive objects that can be manipulated, but are intimately entangled with one another in relationships that have been historically determined and cannot be repeated without repeating that history. The computer gates, on the other hand, are relatively indifferent to their relationships with one another and will respond to manipulation. The relationships in the brain were internally, symbiotically motivated and thus can never be duplicated. The computer gates were never in symbiosis, were internally indifferent to their relationships with one another and thus can always be coordinated to be that way again.² What I am maintaining is that the physical character of the brain processes is an essential ingredient to the fact that the brain is conscious. It is not enough merely to mimic their description at some abstract level *even if that mimicry is enough to reproduce its intelligence. That is, consciousness is a physical specification and not merely functional.* The obvious, and I think desirable, result of this is that consciousness remains a token. I cannot be abstracted from myself.

An objection may be raised by way of possible worlds here. What if on some world a robot with a digital brain arose spontaneously? I do not take this objection seriously because the odds of such an occurrence are so close

² It may be pointed out that I am simplifying the process of making a computer chip. There is indeed a great deal of spontaneous organization of the materials involved in manufacturing a computer chip, and the manufacturing lab can be quite like a huge anti-septic womb. However, this does not change my argument as I am not talking merely about the process of forming the chip but also the end result. The end result does not share the spontaneity found in the formation. It is amusing, and a bit poignant, to imagine that for a brief instant in the manufacturing of each chip a simple proto-consciousness flickers into existence and then fades out again.

to zero even God cannot tell the difference. It is a bit like asking what would happen to our ideas about gravity if an apple fell upwards. Until an example is found we need not take account of it in our theory. These examples are meant only as an imaginative guide to the intuition behind the conclusion being defended here. They are, I realize, controversial. Certainly, programs, abstractly described, may one day be able to simulate the kind of symbiosis that gives rise to consciousness, but as long as it is mere abstract description and not an actual symbiosis, there is nothing ontological at stake for any of the parts and the computer will not be conscious. Computers can only meet the second criteria for consciousness, the organizational criteria. They do not seem to be systems that are governed by the first criteria of the Ontological Principle which combines Meaning to create ontologically individual Things.

The essential difference between a computer program and a conscious mind is the difference between the idea of a tetrahedron and a tetrahedral molecule. While it is true that all tetrahedral molecules are tetrahedrons, it is not true that all tetrahedrons are tetrahedral molecules. Tetrahedrons, like computer programs, are abstract descriptions, while tetrahedral molecules, like minds, are physical phenomena. The physical phenomena may (or may not) be a member of the abstract kind, but the kind in no way exhausts the requirements for being that phenomena. Included in such a discrimination must be certain *specific natural* properties as well as abstract description. Holding that computers have natural properties is not sufficient because they do not have the specified natural properties. Saying that a digital computer can be conscious might very well be like opening the door to every panpsychic nightmare a 1980's New Ager would want to shake a crystal at!

The confusion between conscious minds and computer programs rests on the fact that both computers and brains are physical phenomena, and they both trace out patterns through time. The point to understand, and it is crucial, is that the patterns that they trace out are the results of entirely different kinds of communication, communication which makes brains ontological individuals but computers only sensory individuals. In the case of a brain, the communication is one of Meaning and Receptiveness between ever more complex Things, while in a digital computer it is merely the result of a long and interlocked series of happy coincidences between simple things ? A happens to be closed so B opens, B happens to be open so C closes, C happens to be closed so D closes, etc. Computers are like dominoes falling. Patterns traced out are circumstantial and only incidentally causal. The brain is a chemical reaction in which many different kinds of matter are intimate. The brain is both circumstantial and intrinsically causal.

Finally, I would like to point out that it is not unusual for a physical phenomena to rely necessarily on a combination of both form and substance.

For instance, red is not simply an abstract wavelength, nor is it an expression of substance, being a photon. It is an expression of being a photon of a certain wavelength. Again, let me reiterate that I am not saying that the question of whether or not a computer can be conscious is an empirical question. I am saying that, in principle, digital computers cannot be. Since consciousness is not observable it will always be a question of principle that can only be resolved by appeals to principles of consciousness, whether they are the principles put forth here or another more final and appealing idea.

"So any notion of totality based on a fixed and permanent distinction between thought and reality must collapse when applied to the totality"
? David Bohm

In conclusion, I would like to apply a final gloss to what has been said here. You may have difficulty seeing how the approach taken is new. In a sense, it is not. It can be found sometimes in novels, sometimes in philosophy, sometimes in popular expositions of science. The novelty is merely in suggesting that, refined, it can make a viable approach to theorizing. Nothing discussed defies objective description. Symbiotic systems could be re-termed as feedback systems. Knowledge could be described in terms of the equations of physics. Meaning could be couched in the language of information theory. All consciousness corresponds to a movement in the world, and all movement in the world corresponds to, at the very least, Receptiveness, the building block of consciousness. The only thing gained here is the loss of precision, a loss which is a gain because it results from a shift in perspective. After all, science is in the business of rebuilding the world from an assumed perspective, so it is rather important that the perspective assumed is, from the very start, an inclusive one. The language here is not objective, but empathetic, "inside out" so to speak. As such, it falls prey to the vagaries of interpretation, but it has the strength of allowing us to see how we fit into the world. Consciousness should be a natural and essential part of order, not a mystery that removes us from the world.

Yet the language is, I think, definite enough to reason with, although the reasoning it is amenable to is a different kind of reasoning. Appropriately, it is a symbiosis of analysis and narrative. Our judgments of its truth will rest not only on the rigor of argument and empirical testing, but the beauty of themes played out. For, in the end, consciousness is an expression of Being, and as such can only truly be understood in a process of looking directly at, not abstracting from. This, inevitably, relies not only

on the validity of argument, but on the strength of eye. Scientific analysis and experiment must combine with the methods of literary interpretation, reconcile, and work together so we can see in the mind not only the structure of things but the meaningfulness of what is. Thus mankind comes full circle. St. Anselm said, "Poetry is the cradle of philosophy." Philosophy became the cradle of science. Now science must return to poetry. Fractured perspectives can, after all, come together.

Bibliography

Bohm, David (1980). *Wholeness and the Implicate Order*. Routledge, New York.

Campbell, Joseph (1973). *The Hero With A Thousand Faces*. Princeton University Press, Princeton, New Jersey.

Churchland, Patricia (1989). *Neurophilosophy*. The MIT Press, Cambridge, Massachusetts.

Fodor, Jerry A. (1981). *Representations. Philosophical Essays on the Foundations of Cognitive Science*. The MIT Press, Cambridge, Massachusetts.

Minsky, Marvin (1986). *The Society of Mind*. Simon & Schuster, New York.

Nagel, Thomas. (1974) "What Is It Like To Be A Bat?", *Philosophical Review*, No. 83, 1974, 435-450.

Nagel, Thomas. (1986) *The View From Nowhere*. Oxford University Press, New York, New York.

Nemetz, Anthony (1964). "Some Comments On The Two Cultures." *Proceedings Of The American Catholic Philosophical Association; The Catholic University Of America*, Washington, D.C..

Owens, Virginia Stem (1983). *And The Trees Clap Their Hands*. Wm. B. Eerdmans Pub., Grand Rapids, Michigan.

Penrose, Roger (1989). *The Emperors New Mind*. Oxford University Press, New York.

Prigogine, Ilya; Stengers, Isabelle (1984). *Order Out of Chaos*. Bantam, London.

Searle, John (1980). 'Minds, Brains, and Programs', *Behavioral and Brain Sciences*, Vol. 3, No. 3, pp 417-458.

Tulving, E. (1985) "Memory and Consciousness." *Canadian Psychologist*, No. 26, 1-11.

Whitehead, Alfred North. *Process and Reality (corrected edition)*. Griffin, Donald Ray; Sherburne, Donald W. eds. The Free Press, New York, 1978.

This document was created with Win2PDF available at <http://www.daneprairie.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.