WHAT IS CONSCIOUSNESS?

© May 2017 by Dan Bruiger
dbruiger [at] telus [dot] net

Abstract: Consciousness is a virtual-reality display, created by an internal executive agent, which human beings experience as the self. This display serves to monitor and manage the state of the organism in relation to the world. Conscious experience serves a different function than unconscious processing. It is therefore not superfluous but has limited causal powers within the organism: the causality of agency, rather than the passive (efficient) causality of physics.

1. The modern challenge

The scientific question of the nature of consciousness is relatively modern, a product of the European Renaissance and a budding materialistic philosophy. Before that, and in most parts of the world, the vision of reality was more spiritual and idealist. The material world was but a temporary venue for the soul's journey and no basis on which to explain consciousness, which needed no explaining since it was taken as self-evident or primordial. Rather, events in the material world were explained in terms of agents that were more often spiritual than material. In contrast, in the scientific view of reality consciousness is an incidental byproduct of physical processes, which are considered fundamental if not primordial. A troubling corollary of this view is that consciousness can seem to play no causal role in the life of organisms, even human beings.

The view that will be presented here is that consciousness *does* play a significant causal role. The causality operating within organisms, however, must be understood differently than the "efficient" causality conventionally understood by physics to operate in the non-living world. This is because organisms (and subsystems within them) are *agents* as well as passive objects consisting of "inert" matter. Consciousness is a virtual representation, to itself and for its own purposes, by an executive agent (the conscious self) tasked with monitoring the state of the organism and its environment, planning future action, and coordinating various sub-agencies. Consciousness is thus not useless or "epiphenomenal," but serves a control function for higher organisms that is distinct from that of non-conscious processing.

The challenge for any theory of consciousness compatible with science is to show how, if at all, consciousness can fit within the materialistic framework. Conceivably it might *not* fit, either because it is not a material phenomenon (or not produced by the brain), or because the scientific framework itself is too constricting. Leaving aside religious and idealist views—and considerable evidence for the independence of mind from brain—there have been dozens of approaches to the problem of consciousness within the materialist framework, with perhaps hundreds of variants. Yet, in a sense,

materialism (or physicalism, as it is now called) is as much the problem as it is the solution, since our cultural faith that the world is fundamentally physical leaves little place for subjectivity and consciousness [Crane & Patterson 2000, p1], let alone for an agent such as the soul.

However we conceive of consciousness, human traditions include the assumption that it is the normal and basal state of human being. Yet, since Freud, it has been established that much of human cognition takes place outside conscious awareness. Thus, a current philosophical conundrum concerns whether *all* of cognition could take place unconsciously: whether creatures identical to human beings—except lacking consciousness—are a realistic possibility. A complementary question is whether machines could be conscious, and under what circumstances. An important issue, therefore, is what functional role consciousness plays—in the life of a creature and in the scheme of biological evolution. What purpose, if any, does it serve?

However it has been framed over the ages, at root of the challenge posed by consciousness lies the ambiguity between supposedly real things and subjective experiences *of* them. The very notion of consciousness is ambiguous if only because, like many other psychological terms, it can be interpreted in a phenomenal or in a behavioral sense. It can also refer, for example, alternatively to sentience, to awareness, or to self-consciousness. Such ambiguities hinder even agreement about the nature of the problem. They arguably inhere in reflexive consciousness itself, in the subject-object split. Dualism is not merely a product of language, of European cultural history, or of scientific or philosophic tradition. The very nature of self-reflective mind potentially divides thought about itself, making the problem of consciousness doubly hard. This is an unavoidable aspect of being human, the only escape from which may be to ignore it.

A second ambiguity concerns the nature of the material world. While organisms have long appeared to be agents capable of originating their own behavior, the deterministic view of their material nature led to reductive strategies to assimilate the apparent teleology of self-organizing systems to the standard efficient causality of physics. Behaviorism was relatively successful at accounting for simple behavior in mechanistic terms, but at the strict price of ignoring consciousness. The advent of computers broadened the mechanistic view to enable consideration of more sophisticated behaviors and even the question of consciousness. Above all, it introduced a view of physical processes compatible with mental processes. The computer, like the brain, was a physical system that expresses thoughts.

2. The computer metaphor

Modern theories of mind draw upon computer analogies—essentially, so to speak, to put the programmer in the brain's shoes. This is on a fruitful track, as far as it goes, since it positions the scientist to consider the organism's proper agency, which tends to be overlooked in the physicalist framework. But this is hardly to say that the brain, let alone the organism, is a digital computer. Nor must it mean endorsing a naively realist perspective that presumes cognitive access to the world as it "really" exists. The theoretical constructs of physics (including the philosophy of mechanism) are no less mind-dependent than folk categories based on everyday experience. One might prefer to

think that the entities of the scientific world are what truly exist and that the phenomenal world is a kind of biologically pragmatic illusion. But one could justifiably hold, to the contrary, that the manifest world given in everyday experience is what is real and that the entities of modern science are but pragmatic fictions, "symbolic tools" to help us navigate the world [Sellars 1963, p31]. A third possibility is to consider both on a similar footing: the image of the world given in normal perception *and* the scientific image of the world are alike useful constructions. Perhaps one reason why the scientific study of consciousness was delayed for so long is that the question of what consciousness is and the question of what reality is are awkwardly entangled.

The computational paradigm has been rightly criticized for presenting the brain as disembodied and self-contained—a deterministic structure lacking its own powers of agency. The notion of internal symbolic representation has been criticized for simply regressing the problem of perception to an inner observer. And any approach whatever must eventually come to terms with the paradox of access to the world-in-itself, first addressed by Kant. I will simply assume here that some form of inner symbolic representation (if only memory) is necessary (if not sufficient) to explain both conscious experience *and* the behavior of higher organisms. This assumption is grounded in the fundamental principle that all cognition is co-determined by both the organism and its world. Disregard of this fundamental principle leads to various biases and even absurdities, as some investigators side more with external and some more with internal factors.[1] The very idea of objective reality implies focus "out there," on the object of thought or perception, without regard for the subject. Similarly, focus "in here" on phenomenal experience may disregard the object, leading to pure idealism or solipsism. Extremism is thus built into thought about consciousness. Yet, it is common sense that all forms of cognition—including science itself—involve the interaction of subject with object, however those are conceived. Behavior and experience are alike co-determined by the organism and by the world conjointly. Organism and world also affect each other objectively in various ways. The organism has the capacity to physically shape its environment, as well to define what is salient within its sensory input and respond appropriately.[2]

Because of such subtleties, mind remains the elephant in the room for science; the diverse approaches to it suggest the story of the blind men who each describe a different part of the creature and can reach no consensus. While opinions differ on the relative significance of key factors, a reasonable theory of mind must hold that organism and environment contribute interactively to conscious experience. I will propose that what is missing to comprehend the mystery of conscious experience is a concept of internal agency and communication that accommodates the point of view of the organism itself as the co-creator of both its experience and its behavior.

---

[1] This is perhaps understandable, since the situation resembles an equation in two variables, for which there can be no solution without a second "equation"—failing which (as it must), the only recourse is to arbitrarily hold one "variable" constant to explore the other. Scientific experiments, for example, are designed to artificially isolate a single causal factor.

[2] It adapts the world to itself and itself to the world, which are distinct capacities. One is the result of the organism's motor action upon the world; the other is a result of its cognitive action, which may include information gained through motor action but which affects primarily the organism itself and only indirectly the world.

3. Secondary qualities

In science, the so-called secondary qualities are banished from physical description on the grounds that they are subjective artifacts of the organism's perceptual processing. Consciousness is thereby thrown out with the bathwater of subjectivity. This policy is misleading, to say the least, if *all* cognition, including science, is a mental artifact.

One reason why consciousness is hard to explain scientifically is the tradition in which "secondary" qualities are artificially separated from "primary" ones, effectively on the basis of special properties of the visual sense. Vision seems to transmit a truer image of the world than other senses too closely dependent on the body, which are not considered a reliable source of knowledge. The primary qualities presented in vision (size, location, motion, shape, etc.,) are supposed to represent how nature is carved at its real joints. However, someone is required to do the carving, so there remains a subjective factor at the core of the ideal of objectivity. Moreover, unlike physics concepts based on spatial extension, concepts of mass and force refer not only to vision but also to touch and proprioception. Even the distance senses are body-centric, detect differences only, and serve survival. In contrast, scientific instrumentation measures independently of the organism and thus seems more objective. However, scientific instruments have their own specific properties and limitations as artificial senses. They cannot guarantee objectivity in any absolute sense, since they are extensions of human perception and thought, which in turn is grounded in sensory experience. Even the scientist's experience of reading an instrument is necessarily a first-personal experience [Boltuc & Boltuc 2007, p27]. Yet, the factual knowledge that derives from experience and dominates science is third-personal.[3]

Facts are distilled from phenomenal experience—from the visual sense in particular—but the problem posed by consciousness is that the qualities of experience cannot be reconstituted from such facts alone.[4] The mining of facts (propositions) from phenomenal experience is what gives us a communicable notion of truth. But this extraction seems to be irreversible and comes at the price that the experience itself cannot be reconstituted from such facts. To know what it is *like* to taste a pineapple is not the same as knowing *properties* of the pineapple—or even properties of the whole system including the subject's physiology and brain. Factual information derives from sensory experience; and, as we shall see, experience expresses factual information in an integrated format. The process of that integration is unconscious, however, so we have no first-person access to the information that the brain uses to constitute experience.

---

[3] Propositions are asserted *by* an agent, and in that sense are first-personal statements. 'I see that this tomato is red' is a different proposition, however, than 'This tomato is red'. The first claims responsibility for the perception, while the second does not—which is what by convention makes it "third-personal."

[4] NaCl, for example, is a theoretical abstraction distilled from diverse sensory percepts, which cannot be reconstituted solely from it. If we wish to explain the salty taste and even appearance of NaCl, we must consider also the physiology and intentionality of the perceiving organism. Cf. [Revonsuo 2010, p41]: "In NaCl (common salt), there is nothing constituting the 'salty' quality that we taste in our phenomenal consciousness."

4. The problem of cognitive domains

Another difficulty besetting the study of consciousness is what I call the 'problem of cognitive domains'. Attempts to explain how the mind builds its picture of the external world usually begin with the very picture of the world they attempt to explain. A *domain* is generally a set of elements upon which some operation is to be performed, such as a mathematical function or mapping. The problem of cognitive domains is the circularity that arises when the domain that is the output of a cognitive process is recycled as its own input. This occurs, for instance, when the physical world that appears in conscious experience is presupposed as a point of departure in order to explain its own appearance in conscious experience. The output of mental processing is recycled as the input. An example is how three-dimensional (depth) perception is supposed to be reconstructed by the brain on the basis of information that is projected onto two-dimensional sensory surfaces. These latter, however, are already presumed to exist in a three-dimensional world. How can three-dimensionality be both a mental construction and the real-world basis from which the construction is made? Indeed, how can (subjective) mind provide access to (objective) reality?

This general epistemic dilemma applies not only to ordinary perception, but also to every form of cognition, including science itself. The scientific strategy in regard to the problem of consciousness is to explain the contents of consciousness in terms that originate in the world as conceived by physics; however *that* world is but another content of consciousness originating through the very processes to be explained. If human cognition is merely one possible form and not a god's-eye view, then it would be a mistake to take it as a standard against which to measure the cognition of other creatures.

The peculiar nature of the epistemic situation of any organism is that it must deal with the external world in order to survive; yet it has perceptual access to that world only by means of its own internal processes (which are presumed by a human observer to be processes occurring in the world as conceived by humans). There is no "direct" access to the "world-in-itself" that bypasses cognition. Human perception may bypass or transcend another creature's perception, but it cannot bypass or transcend its own except through scientific instrumentation that is an extension of its perception and cognitive biases.

The human brain is sealed within the skull, with external senses that do not provide open portals on the world but only a stream of data to which it must respond in a manner that allows the body to act in such a way that it survives. The situation is no different if the stream of data is supplied by instruments rather than biological senses. One can assume, for convenience, that the brain's model of the external world must be realistic, but the brain cannot get outside the skull to assess the degree of fit that perception bears to the world [Sterpetti 2016, sec 1]. The only feasible measure of this fit is whether it favors survival, which is a statistical question that can only be known in retrospect.

One could say that physics is our best shot at describing that absolute domain behind the world of appearances, but even it is not exempt from this fundamental dilemma. Science provides conceptual alternatives to our perceptual models, but they are still only theoretical models. While it shifts the stream of data to instrumentation, the

burden of interpretation remains. There is still no access to how the world "really" is "in itself," apart from interactions with instruments and speculation. We may hold that scientific cognition is more accurate than ordinary cognition, where accuracy means the ability to measure and predict; but the ultimate test for both remains viability. Despite the proliferation of human civilization on the planet, science is too recent on the evolutionary scene to conclude that it ultimately enhances our prospects.

If the organism has sense organs, it responds to changes in these through activity that restores a preferred state. (Even without dedicated sense organs, it may respond to changes of its own chemistry, for example, in ways that prove adaptive or not.) Either way, the organism is effectively "flying by instrument" [Oatley (1978)] [Maturana (1978)]. The situation of the brain isolated inside the skull may be likened to navigating a submarine. Another apt metaphor is a flight simulator that happens to represent and control a real plane in flight, in real time. Like the chamber of the skull, the submarine or the simulator is a windowless cockpit, with electrical input from various sensors (other than direct video), so that the pilot is navigating entirely by instrument. The "picture" that emerges in the pilot's mind is a matter of inference from the readings. Of course, such metaphors are limited. For, this "pilot" has never been outside the cockpit or submarine to experience the external world directly. The brain is in the same situation, which might be compared to that of congenitally blind people who sometimes claim to "see" what is before them. The pilot must accept the "picture" the simulator presents as a "true" representation—so long as it doesn't result in destruction. The problem posed by consciousness is to understand exactly how picturing can arise.

The single cell and the human body face the same challenge: to respond to disturbances of its state in ways that favor its continued existence. From an observer's point of view, there is a world external to the cell wall, the skin, or the skull. The challenge for the organism, however, is simply to survive and reproduce, which may or may not entail activities that the observer interprets as modeling an external world. This seems straightforward enough until the organism in question is the human observer herself. There is then a conflict between considering consciousness a true depiction of a real external world or merely an adaptive strategy.

A further obstacle to understanding consciousness is the very *transparency* of perception. We see the world *through* our perceptual processes, which themselves remain invisible to us. This is not a logical but a biological condition. It is no coincidence, but is highly functional. For, the nervous system is organized precisely to provide this illusion of a transparent window on the world. While flying by instrument, the pilot nevertheless navigates and "sees" in air space. (Indeed, modern aircraft are designed to allow the pilot's attention to dwell in air space rather than on instrumentation.) In general the organism has no need to access its internal processing, and doing so in many cases would be counterproductive. As a consequence, the qualitative states of experience have a gestalt and ineffable aspect. It also means that we do not normally question our perceptions, a fact that has social consequences and is the basis of naïve realism. In compensation, it motivates the ideal of objectivity and a functional wariness toward overconfidence in our perceptions.

5. Epiphenomenalism

In the scientific worldview, consciousness must be a byproduct of physical processes within the organism. Indeed, conscious experience seems to occur slightly after the neural events that underlie it. This may seem to leave consciousness without a causal role to play. However, this dilemma of "epiphenomenalism" results from a simplistic and mechanistic view of single causes linearly preceding effects in time. It implies that the conscious self is simply "along for the ride," playing no active role but passively experiencing effects that have already been determined by events to which there is no conscious access. Yet, organisms involve circular and multiple causation, operating on many levels. Let us propose, therefore, that the role of consciousness is precisely to respond, *after the fact*, to events that have already taken place in the nervous system. This hardly negates a role for the conscious self to play, any more than the operations in a computer negate a role for the user to play. Quite the contrary, consciousness may serve as a user interface for making high-level decisions that *do* have causal effects on the operations of the system.

Thomas Huxley's [1874, p240] famous train whistle metaphor does not demonstrate epiphenomenalism nor aptly represent the situation of consciousness. For, only when the system concerned is the engine alone does the sound of the whistle appear to be purely effect and never cause. The whistle *does* have causal effect when the system includes the people for whom the whistle is intended as a signal. The whistle is functional in the overall operation of the rail system. The same is true of consciousness. It may seem to be the passive byproduct of brain processes that would produce the relevant behavior without it. But, in fact, some behavior can only occur consciously, reflecting a high-level causal role in the overall management of the body. While this does not *explain* consciousness, it refutes that it is superfluous.

6. Virtual reality

The strategy to understand cognition in computational terms began with Descartes,[5] the inventor of analytic geometry, whose key intuition was that propositional forms of representation (equations) can generate graphic forms. Early computer simulations of visual cognition involved interpreting visual scenes in terms of internal models, expressed as computer code, which provided a hopeful means to understand the brain's interpretive processes. However, the computer is more equivalent to the disembodied brain in a vat of science fiction lore than to a brain serving an active living body. Moreover, the approach to scene analysis was based on modeling cognition literally as a logical, language-based rational process. These limitations led to approaches that went to the other extreme, emphasizing physical interaction with an environment and discounting the role of internal modeling in favor of using the environment itself as a store of information. Certainly, the external world does serve as an available reference for continually updating internal models, and the organism is clearly an embodied agent; but

[5] Descartes' great realization was that no literal likeness of the external world is transmitted within the nervous system, but only a symbolic representation [Huxley 1874, p210]. This is ironic given the association of his name with a "theater" of inner images.

without some form of internal modeling to update, there could be no interpretative process to make sense of sensory input. In some cases there may be no *need* for interpretation or modeling, when the organism can survive just by managing input-output relations, as in the case of reflexes. This might be so for very simple creatures, but it is no basis on which to understand conscious experience or the behavior of higher organisms. Using the world as its own best representation seems reasonable as a path to action that requires no intervening consciousness. It may work well to explain insect or robot behavior [Dreyfus 2007]. However, beyond explaining *behavior*, the whole import and promise of internal representation is ultimately to explain *consciousness*—how it is possible at all for there to be such a thing as experience. Insects, and even rabbits, may not require internal representation, but humans certainly do.

The brilliant Leibniz could not conceive how the brain as a mechanism could give rise to subjective experience, and this remains challenging to this day. Partly at fault may be our simplistic idealizations of matter, on the one hand, and romantic notions of our idealized being as subjects, on the other. The mechanical models with which Leibniz and Descartes were familiar were hopelessly simple. A modern computer is unfathomably more complex than a pendulum clock, yet is nowhere nearly as complex as a single-celled organism. While it is possible to imagine, for example, *that* an artificial system does all the same information processing as a living organism, we may not be able to imagine the complex processes whereby it does this *in sufficient detail* to properly assert their functional equivalence. Part of the problem may also be that we still intuitively conceive of machines as material (and therefore essentially passive) while we conceive of minds as immaterial (though active). An understanding of computers as *virtual* machines may help to overcome this dualism. For, a virtual machine is primarily a *logical* system even if implemented in a physical system. As such it is immaterial, and so may serve as a bridge across the explanatory gap between mind and matter. One step closer, the task is then to show how a virtual machine can have conscious experience and a point of view of its own.

A computer monitor and keyboard provide an interface between the conscious user and the non-conscious machine. On the computer metaphor, consciousness provides an interface between a particular agent (the self) and pre-conscious levels of processing. The conscious self may register perceptual events that have already been deterministically caused in the nervous system. This is its purpose and does not negate an active role for it to play, any more than the hidden operations in a central processor negate a role for the human user who interacts with them. Quite the contrary, conscious experience is a sort of display for facilitating high-level decisions that may then have a downward causal effect on the operation of the system. For example, we think of a consciously experienced intention, such as desire or willing, as causing an action that follows. More accurately, it represents (to the self) an aspect of the *organism's* intention.[6] It is one phase of an overall action, among others that remain inaccessible to consciousness.

---

[6] Cf. Thomas Huxley [ibid, p244]: "It seems to me that… our mental conditions are simply the symbols in consciousness of the changes which takes place automatically in the organism; and that, to take an extreme illustration, the feeling we call volition is not the cause of a voluntary act, but the symbol of that state of the brain which is the immediate cause of that act." My point is that such experiences as volition are both symbol *and* cause. For, if an agent then acts on the basis of the symbol, the latter is the cause of further action.

The notion that consciousness is a natural kind of virtual reality, an internal simulation of the external world, is a logical extension of the computational metaphor. It is heuristically fruitful, but its limitations as a metaphor must be kept in mind. Literal VR is an *external* input like any other sensory input, enjoyed by people with real brains and bodies already engaged in perceiving external reality. Just as computation has served as a metaphor to understand the workings of the brain leading to *behavior*, so virtual reality provides a metaphor to grasp how the brain can produce the spectacle we call *experience*. Yet it would miss the point to think of "natural" virtual reality as a sort of "goggles" that the brain puts on to have an experience, as the human user does.[7] Or, to imagine the vista presented by the VR as though dancing inside the skull like a holographic image or a miniature replica of the external world [Lehar 2003, sec10]. Unlike the brain, the user of the literal VR device has her own eyes and brain, by means of which to perceive the input from the device. The very point of the metaphor is that the brain is at once both the device that generates the virtual reality and its user. The brain makes use of the virtual reality it generates for its own purposes and benefit, which is generally not entertainment but survival. In this unique situation, the "simulation" cannot be compared to the reality it simulates, because the brain cannot get outside the skull to make the comparison. So, the question of verisimilitude can only be measured (after the fact) by the evolutionary success the simulation facilitates. One may think of this simulation either as reality or illusion, but nothing is accessible outside it to set a standard of comparison by which to decide—except the truths of science, which are also part of the virtual reality and which must ultimately also be measured by evolutionary success.

Perhaps one reason why virtual reality is engaging—and its coarse-graining acceptable as entertainment—is because sensory experience itself is normally so impressionistic. When attending to a visual scene, for example, it is the overall impression and the feeling it evokes that constitute the experience, not a comprehensive survey of all detail. It is an illusion that we fully and uniformly *see* all that is before the eyes; what we "see" is what we notice—in some cases simply *that* there is detail, without putting a finer point on it. To the extent that a virtual reality can present fine enough detail to give that order of impression, it can pass as a substitute for sensory experience.[8]

We understand how a virtual machine can produce a virtual reality for a human user—because we have constructed systems to do this. The challenge is now to grasp how and why the virtual machine would produce a virtual reality *for itself*. How does there come to be an actual panorama from its point of view in space and time, "something it is like" from that viewpoint for the system itself? Effectively, the task is to write a virtual reality program for the use of the system itself rather than for the human user. While I will not attempt to do this, I do propose that the brain creates the "virtual reality" of conscious experience as a narrative to itself, as an efficient way for an executive function to monitor the world, the body, and their relationship, in the service of control.

---

[7] This spectacle does not entail a regression of witnesses, only a singular internal agent with special responsibilities. The buck stops with one conscious witness—the self—which, like the body for which it is a proxy, is also part of the virtual reality. The very point of the virtual reality metaphor is to show how the simulation (which is but code in a computer) becomes an experience.

[8] Similarly, an artist's crude brushstrokes can produce what appears from a distance to be a highly detailed scene characterized by continuity—an effect that gave rise to the term 'impressionism'.

The question then becomes, why is *consciousness* required for this sort of monitoring and control? In other words, since many control functions *are* unconscious, why can't the executive function be unconscious? The answer may lie in the very nature of the executive function as high-level decision maker. Sensory information available to the organism is essentially probabilistic and ambiguous. As in a corporation or government, sub-agencies passing such information on to the executive must package it for decisive action. (It is one thing to debate possibilities; it is quite another to commit to a definite course of action.) This means, ideally, that whatever the conscious self experiences must be *un*ambiguous even when wrong! The very nature of visual perception, for example, is to identify clearly what is perceived, even if the perception is mistaken. (Consider the Necker cube and other unstable ambivalent figures, which *at a given moment* appear definitely one way or another.) What renders something certain is ultimately no more than a decision that it is so [Frith & Metzinger 2013]. What we consciously see is what the brain has decided is there. Moreover, there should be only one boss who decides; it would not be functional to have multiple seats of consciousness competing in the organism. Sub-agencies might be limited to unconscious processing because they lack crucial resources and complexity required for the executive job. Above all, they do not (re)present the situation with the digital clarity associated with consciousness, which is useful foremost in planning action. On the other hand, conscious processes are slow, their judgment is often mistaken, while often we respond quickly and unconsciously to situations in a way that turns out to be appropriate.

7. Causes and reasons

Physical cause is understood as something that happens within a system but is initiated from outside it, and which changes the state of the system (the effect). The cause of this change may be an agent outside the system—in some cases an experimenter. Yet, cause is normally considered an event in a larger system that is also passively undergoing changes, implicitly initiated from outside it. This expansion of the system's border can be extended indefinitely, but this is at the price of a final reckoning in which agency is deferred to a first cause outside the universe. The bill must eventually be paid. Dealing exclusively with closed systems and efficient causes seems to avoid the dualism of matter and mind, but the eventual necessity of a first cause to set the chain of efficient causes into motion simply pushes the mental outside the boundary of the system studied.[9] This is a logical contradiction that science must ultimately face. The challenge is to understand how a system can cause changes within itself, rather than relying on changes initiated by an agent outside it. That is, how to understand the system itself as an agent.

To explain the arising of conscious experience within a physical system, we must look to its actions within its own sphere of reasons, which means looking from its own point of view as an agent. While it might seem feasible to explain the system's *behavior* purely in terms of efficient causes, to ignore the system's own reasons would severely limit such an understanding. It is not possible to understand either complex behavior or consciousness on such a basis. Conscious experience is not the product of the *physical* system as such, which reacts mechanically to causes, but of the *virtual* system

[9] For the early mechanists the first cause conveniently was God.

corresponding to it, which acts for reasons.

The brain, like a computer, is at once a physical and a logical system.[10] A wiring diagram illustrates at once the logic of an electronic device and its physical connections. Whereas there has seemed to be a contradiction between the self as physical and as spiritual or mental, no contradiction lies in the organism being both a physical and a logical system. Nothing in principle prevents intentional and causal descriptions from referring to the same physical system. An organism makes connections itself (by "fiat"), which may lead to action. Attributions of causal connections, on the other hand, are rather made by an *observer* after the fact. An organism is part of a network of efficient causes perceived by the observer as well as being an agent in its own right. The flow of events in the network can be described either in physical or intentional (logical) terms.

8. Fiat

The present theory is that, by its own acts of fiat, and apart from how an observer views them, an organism makes internal connections that mean something to it, thereby creating a point of view from which the world appears real and external. Guided by its embodied existence in the world, the brain in effect *declares* its experience into being.[11] Accordingly, the observer may view some of these connections not only as the result of causal processes, but as propositions representing elements of the external world. In other words, an organism as a logical system can communicate with itself by means of an internal language; and some of this communication will concern elements that appear in the observer's world. The meaning to itself of these internal communications is analogous to the meaning and imagery that emerge in the act of reading or writing, of speaking or listening to a story or discourse, in which the human communicator's brain translates symbols (written or aural) into mental images, thoughts, and feelings, or vice versa.

The language user intentionally *assigns* meaning to symbols that bear no intrinsic meaning. In much the manner of the human language user, the brain's reading and writing of its own internal "language" of nerve impulses gives rise to its conscious experience. This experience does not seem to happen inside the skull because what it refers to is not inside the skull. The whole point of perception is to navigate an external world, not inside the brain.[12] Consciousness cannot be "located," because it is not a thing like a movie screen. The projective capacity, whereby the world is experienced as external, can exist because the relation between intentional connections and neural connections is *like* the relation between the meaning of what you are now reading and the paper and ink (or electronic pixels) that physically convey it. The ink is on the page, the digits are in the computer; but the story unfolds in another domain.

Just as squiggles on a page come alive as a narrative, and mathematical symbols may represent actual relations between things in the world, conscious experience is how

---

[10] Human agents create the logic circuits of the computer, whereas the organism (or nature at large) creates its own logic circuits.

[11] Cf. the divine fiat: "Let there be light!"

[12] The latter would be an interesting second-order natural ability, for which evolution has not equipped us, perhaps for reasons of economy, but which neurophysiology approaches indirectly.

the brain *represents to itself* its own intentional connections that bear significance for it, especially in terms of a putative external world. Conscious experience only occurs when there is need for such explicit representation, apart from which intentional connections may be unconscious, though nonetheless meaningful to the organism.

The self-evident efulgence of sensory *qualities* (the redness of red, the hurtfulness of pain) arises in much the way that the meanings of words do. Sounds and symbols carry vivid meaning as words through a constructive process. Intelligible meanings emerge through the babble of spoken syllables or the symbols on a written page. Algebraic symbols gain numerical significance by the mathematician's declarative fiat: '*let x stand for such and such…*' Similarly, the qualities of experience express the organism's responses to something meaningful to itself. Pain, for example, represents tissue damage and an associated compelling response. The relation between the conscious experience (whether of pain or of auditory tone, for example) and what it represents (whether tissue damage or longitudinal vibrations) may appear arbitrary and mysterious to someone in the observer role, just as the symbols of a foreign language appear mysterious to one who does not understand it [Grahek 2001, p150]. We do not normally question the reasons for our own brain's internal connections, any more than we question the use of familiar words. It is only from a third-person perspective that habitual connections appear arbitrary, because the observer is then not in the role of being the agent that makes them—even when that agent is one's own brain.

9. The language of the senses

An agent does not have to be conscious, but one cannot understand consciousness without considering the internal agency of the organism, as well as its agency in the world. Both aspects of agency must be understood in the context of the evolutionary and ongoing interactions through which internal symbolic systems acquire reference in the world. The very idea of "reality" derives from effects upon and by the body in the physical world, and refers generally to the power of the environment over the vulnerable organism and its own limited powers over the environment. Cognition is thus embodied by its very nature, and embodiment is more than just physical presence. It is a relationship of an *agent* with the world, whether established through the present interaction of the individual or the interaction of the kind over generations of natural selection. The meaning to the organism of its internal communications refers ultimately to its evolutionary history. The organism makes sense of its relations to the world through a history of interactions that *matter to itself.* Motivation is the source of all meaning, and motivation comes from embodied participation in an evolutionary contest, in which survivors are selected to be motivated to survive. The premises of the embodied mind are not arbitrary, nor programmed from outside like a computer, but inhere as values implied in genetic fitness. The organism's evolutionary and personal histories "program" it and provide the reasons for its reasons. It is not wetware that makes the brain different from a robot in regard to consciousness, but its embodied motivations deriving from

evolutionary history, which give it "original intentionality."[13]

Meanings in the lexicon of the organism's internal symbolic language are established through natural selection. Clearly, pain *must* "hurt" if the creature is to survive; sugars *must* taste "good." Similarly, space *must* have the look of "depth" if visual appearance is to guide distance perception to accommodate movement through it. And the world *must* have a solidly "real" look and feel to it if consciousness is to serve negotiating its dangers and promises with due respect for consequences. In such terms, it is less clear why berries should appear red while grass appears green (rather than, say, vice-versa). However, the very existence of color categories (hue) indicates an evolutionary significance, since they reflect needs of the organism more than properties of light or reflective surfaces.

Evolutionary advantage readily explains color *discrimination*—why things appear to be colored differently—but not why a particular wavelength of light is experienced just so and not otherwise. Perceptual discrimination is a form of behavior, which does not necessarily imply conscious experience. It remains to establish the nature and basis of the link between behavior and experience—and the specific role of consciousness. Many species do not enjoy color perception, which is functional within particular ecological niches. Human beings inherited color perception from tree-dwelling primates who occupied a niche otherwise reserved to birds (who have keen color discrimination). To compete with them, primates evolved color discrimination of a similar order [Humphrey 1976, pp. 95-98]. The diets of Old World primates consist significantly of fruits that are yellow, orange or red [Tsou 2013]. It would make sense for these food items to stand out from a background of foliage, and subjectively we think of red as the "opposite" of green.[14] In the forest context, in any case, the color red serves to alert the creature to something singular—whether a ripe fruit or a poisonous creature that has adopted the color code [Dennett 1991, p385]. But other associations would be possible, such as blood, or colorings related to sexuality, or the red of dusk when some predators hunt. No doubt all sensory qualities involve a network of such (possible charged) associations, for which the quality itself stands summarily [Loorits 2014, citing Crick & Koch].[15]

The question of the "meanings" of sensory qualities, such as specific colors, may be likened to asking why a particular meaning is denoted in the English language by a particular word, written and pronounced its particular way, rather than by some other symbol. For the native language user, the association seems natural and unquestioned, though of course it is logically arbitrary and a product of historical accident. The

---

[13] As an artifact, the robot embodies the intentionality of its creators. This does not preclude the possibility that robots might embody their own intentionality, if they are one day self-designing and self-programming. In that case they would effectively be artificial organisms.

[14] A red source of light produces an afterimage and a color shadow that appear closer to blue than to the green of typical verdure. But, then, some fruits are more blue than red. Yellow fruits are closer to green than red, but yellow may also indicate a less ripe fruit.

[15] Cf. [Humphrey ibid]: "We may presume that colour vision has not evolved to… [perceive] inorganic nature, since rainbows and sunsets have no importance to survival. Nor is it likely to have evolved to see simply the greenness of grass or the redness of raw flesh, since those animals which feed chiefly on grass or on flesh are colour-blind… But the most striking colours of nature, those of flowers and fruits, the plumage of birds, the gaudy fishes of a coral reef, are all 'deliberate' evolutionary creations which have been selected to act as visual signals carrying messages to those who have the eyes to see them."

subjective experience of qualities—in this case color—arises from sensory input in a way analogous to how meaning arises from the sounds or characters of language. The internal language of the organism may be no less arbitrary and accidental in its choice of symbols than English. *Some* symbol must be chosen, and by convention it will inevitably come to seem imbued with the meaning it has been made to convey. So, it is backwards to ask why grass "looks" green rather than red; instead, the brain imbues greenness with the association of verdure. Greenness is the *way* we visually experience the totality of associations related primarily to chlorophyll. Redness is the way we visually experience associations related to behaviorally significant exceptions to the verdant backdrop.

The *sensations* of greenness and redness (unlike the words) are not merely linguistic conventions subject to social change, but conventions of neuro-logical organization, with the enduring force of long genetic precedent. Indeed, the human cognitive system gradually adapts to distorting colored lenses or filters in such a way that experience of verdant foliage, for example, is restored more or less to normal greenness [Neitz et al 2002]. The words of a natural language have relatively transient reference, on the cultural time scale. The meanings of perceptual qualities are more stable, being backed rather by evolutionary history. The sensation of redness is inherently different than the sensation of greenness precisely because of the real-world things it refers to, presently and in our evolutionary history, from which it cannot be arbitrarily dissociated.

Whether or to what extent perception is biologically functional is a different question from whether it corresponds to reality.[16] Realism does not imply resemblance of an internal model to external reality (as with a painting), but conduct that permits survival in a context. What is real for the organism is what can affect its well-being and which constitutes a potential reason for action [Beaton 2009, sec 2.3.3]. Since only the behavioral concomitants of consciousness can enter into the mechanics of selection [Jonas 1966/2001, p127], it might seem that nothing in the process of natural selection sheds light on the adaptiveness of conscious experience itself [Horst 1999, p44]. However, natural selection can account for the organism's motivating *valuations*, which are the basis in common of behavior and experience alike. The organism's responses are based on valuation, the conscious experience of which is *affect*. While behaviors of aversion and attraction need not involve consciousness, consciousness necessarily is grounded in the values behind such behaviors. Feeling, with the judgments behind it, is therefore central to consciousness. But valuation lies behind all intentional behavior, whether conscious or not, and so is necessary for consciousness while not sufficient.

Much of phenomenal experience is not directly related to survival. And much of functional behavior is unconscious. Pain in mammals, however, seems to serve an alerting and motivating function beyond simple reflex; it favors protective behavior toward the injured part. In contrast, insects—which lack specific nociceptors but have pre-programmed avoidance and escape responses—probably do not experience pain or have any need to. In particular, they do not exhibit protective behavior toward damaged limbs or other parts, but stoically attempt to carry on as usual in the face of injury or missing parts [Eiseman et al 1984, p165].

Pain is at once feeling and response—as indissolubly related as subject and object.

---

[16] The observer evaluates functionality in her own terms and may not be able to determine all the functions something serves. Moreover, a feature may serve multiple functions [Dennett 1987/1998, p.319].

Indeed, the feeling includes a reading of the response [Humphrey 2000b, p13]. To the organism itself, pain *means* avoidance, or protection of an injured part, as well as the damage itself. But the pain response has two phases, corresponding to two neural pathways (c-fibers and a-fibers).[17] One is a quick reflex reaction—removal of the hand in response to contact with a hot surface, for example. The slower response of lingering painful sensation reflects an ongoing *internally generated* stimulus, and evaluated acknowledgment of the state of tissue damage. Some version of it may persist throughout the process of healing, which takes place over time; the response associated with it is protective behavior. Just as the quality of an auditory tone emerges from impinging successive wave fronts, we might think of the ongoing nerve signal responsible for lingering pain as consisting of a succession of deliberate re-iterations of the first reflex impulse—a reverberation or reactivation loop persisting in time [Humphrey p204-5]. The integration over time of these reiterations constitutes the *quality* taken to signify the persisting damage—namely, the hurtfulness of the pain. This conscious experience of pain carries several implications: first, that the initial reflex was not sufficient to avoid damage; second, that the injured part must be favored during healing; third, the lesson to the creature to avoid such stimulus in future; and fourth, the lesson to the observer that the experience is not a passive suffering but an active management of the situation.

10. Valuation

Pain and pleasure are responses that obviously involve valuation and response. But what is the valuation involved in sensory qualities that do not manifest aversion or attraction? Can we hope to understand, in a parallel way, qualities that do not seem to involve a reflex response or other associated behavior? How to understand, in terms of valuation and affect, such qualities as color, auditory tone, odor, etc.? A reflex response to a proximal stimulus may have immediate benefit for a creature. However, on the model of pain, it is not this immediate reaction but the subsequent mediated *valuation* that is salient for the specific qualities of sensations and upon which *future* action is to be based. The reaction to a hot surface occurs before the onset of pain. But then there is an ongoing response to tissue damage that is based on ongoing valuation. This response continues to carry with it the implied (but not actualized) reflex motor response.[18] The question is how to interpret, in terms of affect, perceptual qualities that do not seem to imply any action, let alone to involve a reflex impulse.[19] The problem presented by the seeming detachment

[17] As Dennett [1978, p200-202] points out, the physiology of pain is more complicated than this, involving separate channels through the "old brain" and the "new brain," and also the possibility of other pathways influencing the experience of pain. This should not affect the argument here, which concerns the grounding of the quality of pain in the associated response.

[18] There is neurological evidence that phenomenal experience involves efferent as well as afferent nerves [Ellis 2000, p44]: "Consciousness always involves efferent activity, defined as neural activity generated by the organism itself, for purposes of its own survival and well-being, rather than from passive stimulation by incoming sensory signals."

[19] But which may have once in evolutionary history. Cf. Nicholas Humphrey [2000a]: "Now, it is true that, today, these sensory responses are largely internal, covert and private. But, or so at least I want to argue, it was not always so. Rather, these responses began their evolutionary life as full-

and objectivity of the visual sense (and hence, perhaps, that of science) is to understand how it derives so indirectly from subjective affect. It does not seem that there can be a behavior comparable to attraction or avoidance associated with a given wavelength of visible light, for example, which is too weak a stimulus to directly affect the organism through simple contact [Jonas 1966/2001, p29]. (For most wavelengths, there is little impact involved for the cell in an encounter with a photon.) Rather, the organism must *interpret* such a weak stimulus as having a significance requiring *considered* action—which is complex behavior, not simple reaction. One may speculate that a global *image* is built up essentially from primitive *responses* at a lower level, in the way that a digital image is built from pixels [MacLennan 2005]. Each "pixel" may represent a simple valuating judgment (1 or 0), but the image that emerges with scale integrates that information, facilitating response on another level. The minute scale of the individual impulses (pixels) allows both their miniscule energy and their individual import to be absorbed into a larger synthesis from which the subject may be relatively detached.

All nervous activity consists in the same sort of electrochemical signals. All extero-senses derive from the cell membrane or skin of the organism. The senses differentiate in various ways from that tissue, so that light, for example, is no longer (just) a proximal stimulus with its immediate physical effect, but primarily conveys information from a distance, which no longer bears the immediate import of direct contact and which engages a different level of response: monitoring and evaluating from afar. The distance senses are by definition freed from the need for immediate response. Compared to touch, for example, visual and aural *qualities* are accordingly dissociated from such response.

Though sensory qualities are products of detailed neural events and reflect underlying structure, they gloss over and integrate such events and structure, masking their etiology or composition. Through such qualities the organism (or its executive agent) displays to itself certain of the products of unconscious processing. Color or sound experience synthesizes many sources of information about structure in the environment to yield an appraisal that is useful to the organism. This appraisal is the end product of processes bearing little information about the processing itself. Sensation implicitly and indirectly conveys information about properties such as wavelength of light, reflectance, sound frequency, etc.—which are explicitly measured by the scientific observer as *quantities*. But the effects of individual photons of light or wave fronts of sound are synthesized into a seamless emerging *quality*.

Now, I make the general claim that qualities emerge in conscious experience through a process that seems to be a matter of unconscious "filling in" or "completion" of what would appear to a third-person observer as mere structure, fact, or data.[20] The mostly anomalous circumstances in which this process is even noticed include laboratory

---

fledged bodily behaviours that were unambiguously in the public domain — and, what is more, as behaviours with a real adaptive role…"

[20] I use the term *filling-in* advisedly, in view of Daniel Dennett's well-known critique of the notion: "The fundamental flaw in the idea of 'filling in' is that it suggests that the brain is providing something when in fact the brain is ignoring something" [Dennett 1991, p356]. His valid point is to distinguish between representing a continuous phenomenal field and representing *that* it is continuous. It is precisely this operation, of representing *that*, which I call *fiat* and claim as the normal basis of the experience of qualities. Let us note also that such data, facts, or structure were abstracted or extracted from phenomenal experience in the first place.

studies of perceptual completion effects, habituation, perceptual adaptations of various sorts, and phenomena of spatial projection. In these acts of fiat, features of the subject's phenomenal experience typically seem to go beyond the facts or events noted by the observer. That is, the subject's brain "fills in" what is "not really there" to someone in the role of third-person observer. I contend that tones, colors, smells and qualities in general may be understood in terms of such completion or projection effects, while offering the caveat that what is deemed "really there" involves a fundamental bias deriving from the scientific use of the third-person point of view.[21]

11. The role of consciousness

While unconscious mental processing precedes and underlies the conscious experience that depends upon it, this does not render the experience itself superfluous. Rather, the conscious experience indicates explicit recognition of that particular unconscious processing, after the fact, by an executive agent (the conscious self). The hand withdrawing automatically from contact with a hot surface attempts to avoid the original stimulus; the persisting experience of pain after the contact marshals the conscious attention of the organism to protect the injured part and avoid further or future injury. It reminds the organism to favor that part in order to facilitate a healing process that takes time.

A similar divide between quick and slow pathways seems to exist in the visual system too, where an initial fast wave of visual processing happens outside consciousness and is made available to subsystems for immediate reflex responses. This is followed by a slower phase of "recurrent processing" that involves integration of various brain areas leading to conscious experience [Revonsuo 2010, p215, citing Lamme]. Here too sensory qualities emerge from reiterated signals and serve a different purpose than the initial signal.

We have seen how a conscious experience can be a separate event from a reflex response associated with it, serving a different purpose with a different associated behavior. It is now time to acknowledge that conscious experience in general serves a purpose distinct even from the non-conscious processing that underlies it (and which differentiates it from the reflex). Consciousness is a form of monitoring of the outputs of various activities of the nervous system. Conscious attention seems to be required in novel, complex, or otherwise mobilizing situations. If the action cannot be done by rote, if it confronts a novel or demanding situation, or if it requires planning and forethought, then conscious attention comes into play. This suggests that consciousness makes sensory input available to higher centers for planning or dealing with novelty.[22] As a state of the system, one role of consciousness is to muster additional resources to deal with situations that are not already handled by existing automatisms.[23]

Consciousness plays a different functional role than behavioral responses that can

---

[21] For a more detailed account, see [Bruiger 2017, secs 8.1 and 8.2].

[22] It seems to involve a synchronous broadcasting of information globally, especially to parts of the parietal and prefrontal cortices [Kandel 2012, p464-5].

[23] The conscious state is called upon to gain new mastery, to learn improved algorithms—effectively, to restore a state no longer requiring consciousness! [Solms 2014, p179]

occur without it, and presumably involves distinct neural processes. Many motor tasks are executed "automatically," and in some situations one is conscious *of* the action, which nevertheless seems to "do itself." Awareness of initiating voluntary activity comes slightly *after* the neural processes that have actually caused it, yet this awareness serves as the basis for choosing *future* action, or action in a larger context [Frith & Metzinger 2013]. Consciousness is thus is not epiphenomenal. It plays a specific causal role, and is the specific effect of that role being carried out. The agent in this role is known to itself and other agents as the "self" of that individual, whose job entails a first-person perspective [Metzinger 2010, p29]. Though it cannot take full responsibility for the actions of the organism as a whole, it seems to appropriate that identity to itself and is often held accountable by others. Like the CEO of a corporation, it is more than a figurehead but less than an absolute monarch. It is a sort of virtual representative of the organism tasked with specific responsibilities, which include monitoring a real-time virtual-reality version of the external world and managing the organism's relationship to it. Consciousness is also closely bound up with memory [Lockwood 1998, p84]. The monitoring function serves to register information coming into the system, tagging it for future retrieval. The experience of selfhood and will are the human first-person versions of a control function to guide the molar action of the organism. The self's powers may be limited as "captain of the ship", but it is no mere stowaway.[24] Being captain does not mean micromanaging every operation involved in running the ship, only holding a certain executive authority.

12. Summary conclusion

In conclusion, consciousness is a virtual-reality display, created by an internal executive agent for use in monitoring and directing the state of the organism in relation to the world. The display represents information about the world and also about the organism itself and its priorities, in accord with the principle that all cognition is co-determined by organism and environment. Conscious experience serves a different purpose than the unconscious processing that underlies it. It is therefore not superfluous but has limited causal powers within the organism.

While I have answered the question, 'What is consciousness?', perhaps no answer to this question can seem entirely satisfactory. Because the very nature of consciousness is to look out upon a world, it is challenging to imagine that consciousness is anything but an open window on that world. On the other hand, the scientific world held to be objective is hardly the one revealed in ordinary experience. Despite scientific reductionism, there is a fundamental disconnect between the entities of theoretical physics and the objects of everyday experience. Whether or not this cognitive dissonance troubles scientists personally, it reflects the convention to exclude subjective expression from scientific description. Until quite recently, consciousness itself was in exile as a topic for scientific investigation. Indeed, the very concept of subjective experience is but a recent development in human history. For such reasons, and because we *are* the consciousness that inquires after itself, it remains existentially challenging to conceive just what consciousness is.

---

[24] PBS documentary, "The Brain", by David Eagelman

REFERENCES

Beaton, Michael J. S. (2009) "An Analysis of Qualitative Feel as the Introspectible Subjective Aspect of a Space of Reasons" (Doctoral thesis)

Boltuc, Nicholas and Boltuc, Piotr (2007) "Replication of the Hard Problem of Consciousness in AI and Bio-AI: An Early Conceptual Framework" preprint for Anthony Chella & Ricardo Manzotti (eds.), *AI and Consciousness: Theoretical Foundations and Current Approaches*. Aaai Press, Merlo Park, Ca

Bruiger, Dan (2017) "Can Science Explain Consciousness? Toward a Solution to the 'Hard Problem'" www.academia.edu/31984553; www.philpapers.org/rec/DANCSE-2

Crane, Tim and Patterson, Sarah (2000) "Introduction" in *History of the Mind–Body Problem,* edited by Tim Crane and Sarah Patterson, Routledge

Dennett, Daniel C. (1978) *Brainstorms* MIT Press

Dennett, D.C. (1987/1998) *The Intentional Stance* MIT Press

Dennett, D.C. (1991) *Consciousness Explained* Little Brown & Co

Dreyfus, Hubert L. (2007) "Why Heideggerian AI Failed and How Fixing it Would Require Making it More Heideggerian", Sec III, preprint for *Philosophical Psychology* 171(18):1137-1160

Eisemann, C. H.; Jorgensen, W. K.; Merritt, D. J.; Rice, M. J.; Cribb, B. W.; Webb; P. D.; and Zalucki, M. P. (1984) "Do Insects Feel Pain? – A biological view" Experientia 40

Ellis, Ralph (2000)"Efferent Brain Processes and the Enactive Approach to Consciousness" *Journal of Consciousness Studies* 7, no. 4, 2000

Frith, Chris D. and Metzinger, Thomas K. (*2013)* "What's the Use of Consciousness?" *https://www.blogs.uni-mainz.de/fb05philosophie/files/2013/04/Frith_Metzinger_Regret_2016_penultimate.pdf)*

Grahek, Nikola (2001) *Feeling Pain and Being in Pain* Oldenburg: Hanse Institute for Advanced Study, p150, quoted in Tim Crane "Cosmic Hermeneutics vs. Emergence: The Challenge of the Explanatory Gap" http://philpapers.org/archive/CRACHV-2.pdf

Hameroff, Stuart R. et al, eds. (1998) *Toward a Science of Consciousness* MIT Press

Horst, Steven (1999) "Evolutionary Explanation and the Hard Problem of Consciousness" *Journal of Consciousness Studies* 6, No 1, 1999

Humphrey, Nicholas (1976) "The Colour Currency of Nature" First published in *Colour for Architecture*, ed. Tom Porter and Byron Mikellides, pp. 95-98, Studio-Vista, London

Humphrey, Nicholas (2000a) "THE PRIVATIZATION OF SENSATION" In *The Evolution of Cognition*, ed. L. Huber and C. Heyes, pp. 241-252, MIT Press

Humphrey, Nicholas (2000b) *How to Solve the Mind-Body Problem* Imprint Academic

Huxley, Thomas (1874) "On the Hypothesis that Animals Are Automata, and Its History" *Collected Essays I*

Jonas, Hans (1966/2001) *The Phenomenon of Life: toward a philosophical biology* Northwest University Press

Kandel, Eric (2012) *The Age of Insight: the quest to understand the unconscious in art, mind, and brain* Random House

Lehar, Steven (2003) "Gestalt isomorphism and the primacy of subjective conscious experience: A Gestalt Bubble model" *Behavioral and Brain Sciences* 26, p376, (discussing Pessoa et al 1998)

Lockwood, Michael (1998) "The Enigma of Sentience" in Hameroff et al (1998)

Loorits, Kristjan (2014) "Structural qualia: a solution to the hard problem of consciousness" *Front Psychol* 5: 237

MacLennan, B. J. (2005) "Consciousness in Robots: The Hard Problem and Some Less Hard Problems", Sec 2B, preprint of ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication

Maturana, H. R. (1978) "Biology of language: the epistemology of reality". In: Miller, G. A. & Lenneberg, E. (eds.) *Psychology and biology of language and thought: essays in honor of Eric Lenneberg*, Academic Press, New York

Metzinger, T. (2010) "The Self-Model Theory of Subjectivity: A Brief Summary with Examples" *Humana.Mente* – Issue 14 – July 2010

Neitz, J., Carroll, J., Yamauchi, Y., Neitz, M. & Williams, D. R. (2002) "Color Perception Is Mediated by a Plastic Neural Mechanism that Is Adjustable in Adults" *Neuron*, 35, pp783–792

Oatley, K. (1978) *Perceptions and representations: the theoretical bases of brain research*. Methuen, London; Free Press, New York

Revonsuo, Antti (2010) *Consciousness: the science of subjectivity* Psychology Press

Sellars, Wilfred (1963) "Philosophy and the Scientific Image of Man" from Sellars *Empiricism and the Philosophy of Mind* Routledge and Kegan Paul

Solms, Mark (2014) "A neuropsychoanalytical approach to the hard problem of consciousness" offprint from *Journal of Integrative Neuroscience*, Vol. 13, No. 2

Sterpetti, Fabio (2016) "Models, Brains, and Scientific Realism" In L. Magnani & C. Casadio (eds.), *Model Based Reasoning in Science and Technology. Logical, Epistemological, and Cognitive Issues*. Springer 639-661 http://philpapers.org/archive/STEMBA-5.pdf

Tsou, Jonathan Y. (2013) "Origins of the Qualitative Aspects of Consciousness: Evolutionary Answers to Chalmers' Hard Problem" preprint for chapter in Liz Swan (ed.), *Origins of Mind* Springer 259-269