The Quest for Consciousness: A Quantum Neurobiological Approach.

Henry P. Stapp
March 1, 2006

Christof Koch has recently written an extremely accessible book, *The Quest for Consciousness: a Neurobiological Approach*. It summarizes in a clear way the main known biological facts that seem particularly pertinent to the mind-brain problem, along with an attempt to understand these facts within the framework that he has been developing in collaboration with Francis Crick. This theoretical framework, although only one of many being pursued at present, is probably as close to main-line neurobiology as any. This article contrasts Koch's way of accounting for certain time-related psycho-neurobiological data with a physics-based understanding of these data.

By a physics-based understanding I mean an understanding based on the fundamental principles of physics, as they are understood today by physicists. This understanding differs profoundly from the one that prevailed at the end of the nineteenth century. During the first part of the twentieth century it became clear that the earlier approach was conceptually unable to account for the results of the observations that we make upon systems whose behaviors depend sensitively on the motions of their atomic constituents. Physicists, lead by Niels Bohr, Werner Heisenberg, and Wolfgang Pauli, erected the new and empirically successful physical theory based not on the idea of evolving material substances, but rather upon the idea of a sequence of abrupt events. Each of these events is *psychophysical* in the sense that it has, on the one hand, certain aspects that are described in the ordinary language that, in Bohr's words, we use to communicate to others "what we have done and what we have learned", and, on the other hand, other aspects that are described in the mathematical language of physics. These psychophysical events link certain physically described aspects of the systems we observe to events in our minds. These mind-matter connections were subsequently converted by John von Neumann to connections between the minds and the brains of human observer-participants. This introduction, at the level of the basic physics, of events that relate physical descriptions of our brains to events in our streams of consciousness has led many physics who are interested basic ontological issues to the conclusion that twentieth century physics may provide the a more adequate foundation for understanding the connection between our brains and our minds than its nineteenth century predecessor. The latter is, strictly speaking, an approximation to the former that systematically leaves out the physical effects of our conscious choices that twentieth-century physics explicitly brought in.

This possibility that the inclusion of the mind-brain connections specified by basic physical principles might offer a more rationally coherent and useful conception of the connection between our conscious thoughts and their neural correlates than the prevailing classical-physics-based ideas in neuroscience can allow has recently been explored in some detail. (Schwartz, Stapp, and Beauregard 2005, Stapp 2004, 2005). The encouraging findings will not be reviewed here. However, readers of this article do need a sufficient understanding of what is meant here by "orthodox contemporary quantum theory of the mind-brain connection". Hence, before describing the contrasts between Koch's explanations of the psychobiological data and the physics-based account, a brief description needs to be given of the significant contrasts between these two alternative theoretical frameworks.

Physics-based approach to the connection between mind and brain.

It is by now widely appreciated that the materialist-deterministic conception of nature that prevailed during the eighteenth and nineteen centuries fails to account for the macroscopic properties of systems whose behaviors depend sensitively on the behaviors of their atomic-sized constituents. That earlier theory was replaced during the first half of the twentieth century by quantum theory. The basic mathematical change was to replace the *numbers* of the older theory by *operators.* The effect of this replacement is to change the theory from one putatively about intrinsic observer-independent properties of physical systems to one explicitly about the observed feedbacks from actions performed upon an observed system by an observing system. Observing systems that stand apart from observed physical systems, but that act upon them in chosen ways, are thus brought crucially into the dynamics of the combined observer-observed system.

The older theory, classical mechanics, has *one* single dynamical process, the continuous evolution of the physical universe in accordance with Newton's laws, or some classical-type generalization of them. In contrast, the new theory, quantum theory, is built around a sequence of discrete events that is governed by four processes. Each such event is, from an ontological perspective, simultaneously: (1), an actualization of certain mathematically characterized potentialities created by prior events; (2) a creation of potentialities for future events; and (3), an experiential event that constitutes an increment in knowledge.

One of the four processes governing the flow of these events is called Process 1 by von Neumann. It is *a mathematically characterized abrupt action of an observing system upon an observed system*. This action has the effect of specifying a particular question about the observed system. Another process, called Process 2 by von Neumann, is a continuous deterministic process. It is the quantum analog of the single deterministic process of classical physics, but it acts only during the intervals between successive abrupt events. The third process is the abrupt event of answering the question specified by Process 1. It is called by Dirac a "choice of the part of nature". It is *random*: it conforms to specified statistical conditions. This "Process 3" is where the famous "random element" enters into quantum mechanics. The fourth process is called Process 4. This action is what Niels Bohr calls a "free choice on the part of the experimenter." This choice specifies which particular probing action---taken from among a host of allowed possibilities---will be instigated by the experimenter. This choice is not constrained or limited by any currently known law, statistical or otherwise, and is treated, in actual scientific practice, as the conscious choice of a human agent. Because this choice is not fixed by any *known* laws it is, in this very specific sense, a "free choice".

Quantum mechanics is, in this way, explicitly about observations and their outcomes, both of which are represented by abrupt events that suddenly change the physically described properties of the observed system in a way concordant with increments of knowledge of the observer. In the von Neumann formulation the directly observed system becomes the brain of the observer, and the mind-matter connections are converted to mind-brain connections. This re-conceptualization of physics is radically different from the earlier classical conceptualization, which effectively leaves out all experiential realities. The quantum framework is more *complete* than the earlier classical framework, because it incorporates descriptions of both our physically described brains and also associated experiential realities, and it specifies the form of a dynamical connection between them. Both of these aspects of our understanding of nature are essential components of science, and quantum theory is expressly designed to bring usefully into science the fact that "in the great drama of existence we ourselves are both actors and spectators." (Bohr, 1963, p. 15: 1958, p. 81)

An important characteristic of this quantum conceptualization is that the substantive-matter-like aspects, have dropped out. The theory is about: (1) abrupt *events*, each of which is tied to an experiential increment in knowledge; and (2) *potentialities* for such events to occur. Events are not substances, which, by definition, endure. And the potentialities have an "idea-like" character because they are like an "imagined" idea of what the future events *might* be, and they change abruptly when a new event occurs. Thus neither the events nor the potentialities have the ontological character the substantive matter of classical physics. Yet the predictions of quantum mechanics encompass all of the known successes of classical mechanics.

A second important feature of this quantum ontology is that the conscious "free choices" --- which are not determined by any known law---can influence the course of the psychophysical events. The principle of the "causal closure of the physical" is therefore not enforced in orthodox quantum theory, not only because of the entry, via Process 3, of random elements, but also, and more importantly in the present context, because of the entry of physically effective Process 4 conscious free choices. These latter choices can influence the *objective potentialities* for future actual events, and thereby affect also what actually happens.

A third important point is this. By virtue of the quantum rules themselves, a *conscious intent* to perform either a physical action or a mental action (say to recall something or to focus one's attention on something) *can have*, by virtue of a well known quantum effect---the quantum Zeno effect---the physical effect of holding in place a pattern of neurobiological activity that tends to cause this imagined, or expected, or intended experience actually to occur. Such a pattern of brain activity---namely one such that its *persisting existence for a sufficiently long time* tends to cause the (experience of the) contemplated action to occur---is called a *template for action.* How it comes about that conscious free choices can, by virtue of the quantum psychophysical laws, cause a physically described *template for action* to be held in place for an extended period has been described elsewhere (in most detail in Stapp 2004), and will not be reviewed here.

The earlier articles, cited above, described many applications of twentieth-century physics to the data amassed in scientific studies pertaining to volition. The present article is an application to data pertaining to the character of conscious perception.

<u>Koch's The Quest for Consciousness and the NCC's</u>

Kock's title advertises his book as *The Quest for Consciousness*, and he speaks briefly, at the beginning (and also in several other places) about the mystery of consciousness---about the puzzle of *why certain activities in our brains are accompanied by conscious experiences.* But he says (p.xv) that "I argue for a research program whose supreme aim is to discover the neural correlates of consciousness, the NCC. These NCC's are defined (p.xv: p. 341) as:

"The minimal set of neuronal mechanisms or events jointly sufficient for a specific conscious precept or experience."

He immediately adds: "*This is what this book is about."*

Notice that he does not say "necessary and sufficient". This means that he is not suggesting an *equivalenc*e between a conscious experience and its neural correlate. Indeed, he says explicitly (p. 19) "The characters of brain states and of phenomenal states appear too different to be completely reducible to each other." Thus he distinguishes brain activities from streams of consciousness, and his definition allows a conscious experience to occur without the associated NCC occurring. He emphasizes, right from the start, that his book is about the NCC's, not about consciousness per se, and he acknowledges that it does not address the big mystery of consciousness: *why does it exist at all*. In fact, he notes (p.334) that his project of identifying the NCC's is what David Chalmers calls "The Easy Problem". Chalmers argues that solving this East Problem leaves unresolved The Hard Problem of consciousness: "why does it exist at all"; the mystery of why "the causation of behavior should be accompanied by subjective inner life."

Francis Crick, in his Foreward to Koch's book, says simply that "our strategy has been to try *firs*t to find the neural correlates of consciousness." [My emphasis.] But Koch says that "Characterizing the NCC is one of the ultimate scientific challenges of our times." This apparent elevation by Koch of the *first* part of the attack on the problem of consciousness, namely the "Easy" NCC part, from *preliminary* to *ultimate* status is probably connected to his assertion that: "I suspect that the Hard Problem …will disappear once one has solved the Easy Problem." (p. 334).

Can we identify the root of this major difference between Chalmers' position and Koch's? I think so! Chalmers, in his characterization of his position says. "I have not disputed that the physical world is causally closed or that behavior can be explained in physical terms." (Chalmers 1996, p. xiii). But this puts him essentially in the arena of classical physics, where consciousness is epiphenomenal: where conscious can do nothing in the physical world not done by the physical aspects alone. Given Chalmers's underlying acquiescence to epiphenomenalism, it is no wonder he encounters a hard problem.

Koch seemingly adopts a far more reasonable position! He apparently believes that consciousness can have real effects in the physical world. "I argue that consciousness gives access to … planning…Without consciousness you would be worse off." (p. 4) He asks " Why then, from the point of view of evolution, does consciousness exist? What survival value is attached to consciousness?" (p. 2) "In a fiercely competitive world consciousness must give the organism an edge over non-conscious zombies" (p. 231) "Operationally consciousness is needed for nonroutine tasks…(p. 12)." "The belief that phenomenal consciousness is real but impotent to influence events in the physical world continues to be remarkably widespread among modern philosophers. While this idea cannot, at this point, be proved false, it can be undermined… (p. 238).

Thus Koch appears to be saying, in many places, that, according to his thinking, consciousness has real physical effects. Presumably, his position is that consciousness *itself* is doing these things. No one doubts that neural activities, or brain activities, can have physical effects: the issue is only whether consciousness *itself, which he recognizes as being inequivalent to brain state,* can have such effects. If the NCC's, and the entailed brain states, by themselves, were able to do whole job, then there would be no justifiable reason to say that (inequivalent) *consciousness* has important effects, particularly in this context where the connection between consciousness and brain states is a key issue.

If Koch indeed believes that consciousness is *not* epiphenomenal---that it does things *not* done by the NCC's and the physically entailed brain states alone, and hence can have an evolutionary reason to exist---then he is going outside classical physics, which, because it enforces the principle of the causal closure of the physical, cannot allow anything not *equivalent* to some physically describable property to have physical consequences. Thus when Koch claims (p. 11) that his non-epiphenomenal "conscious is fully compatible with the laws of physics" he cannot be referring to the laws of classical physics, because classical physics makes consciousness, insofar as it is not equivalent to a physical property, epiphenomenal. Since he cannot, rationally, be referring to the laws of deterministic classical physics, he must be referring to the laws of quantum physics, which do allow our conscious choices to influence brain activity, without being equivalent to any physically describable brain activity!

Koch's explanations of the psycho-neurobiological data are in terms of the essentially classical idea of "victories of coalitions". Thus his model does not appear to bring in the quantum effects that can actually allow consciousness to possess the non-epiphenomenal status that he ascribes to it.

It must be noted, however, that Koch does not always maintain a sharp distinction between the effects of consciousness and the effects of its NCC's. For example, on page 18 he begins, promisingly, with the clean assertion "As I shall argue in Chapter 14, it is quite unlikely that consciousness is a mere epiphenomenon. Rather consciousness enhances the survival of its carrier." But he then immediately goes on to say "This means that the NCC activity must affect other neurons in some manner." But this is a non sequitur if, as his definition allows, the conscious event is not equivalent to its NCC, and can occur without it.

This basic conceptual problem (namely the epiphenomenal character of consciousness within classical physics) points to the logical need, if one respects the accepted laws of physics, to bring into the model the quantum effects that allow consciousness to be non-epiphenomenal. The interesting question, then, is whether this shift, demanded by rationality, has benefits beyond its mere rationality.

Differences between Koch's explanation of some neurobiological data and the physics-based explanation

Buds of Perception

One of the chief features of (conscious) perception is its discrete "all or nothing at all" character. William James [1911] said of percepts that: "your acquaintance with reality grows literally by buds or drops of perception. Intellectually and on reflection you can divide them into components, but as immediately given they come totally or not at all."

In view of the *continuous* character of the evolution of the state of a classically conceived brain it is puzzling, a priori, from a classical-physics standpoint, why perceptions should have this all-or-nothing-at-all character.

Koch's FIGURE 15.1B pertains to an experiment in which a subject observers a light source that is abruptly turned on, then maintained at a steady intensity for a certain interval, and then abruptly turned off. The figure shows a variable that Koch identifies as a "the 'critical' activity at the essential node for brightness". This variable rises with time until, at time $T_{ON,}$ it reaches a conjectured threshold, Threshold_{on}. This variable then becomes an NCC, and it remains so until, at time $T_{OFF,}$ it drops below an OFF threshold, Threshold_{off}. During the interval between T_{ON} and T_{OFF} this critical amplitude, labeled NCC, varies greatly, but the brightness of the light is experienced as *constant*. But why is the experienced brightness zero before $T_{ON,}$ then *constant* over the period between T_{ON} and $T_{OFF,}$ while the strength of the driving amplitude is greatly varying, and then abruptly zero again? One might expect, a priori, from classical physics, that if the *presence* of the experience depends on the strength of an amplitude then the *quality* of the experience should also depend upon the strength of that amplitude, and should therefore vary over a period during which the strength of driving amplitude is varying greatly. Why is this not the case?

Koch's explanation of the empirical constancy is this. Coalitions of neurons, firing in coordinated ways, battle for supremacy. When some coalition reaches "some sort of threshold" (p. 253) it brings about coordinated firings in many parts of the brain "until a stable equilibrium is reached." At this point an experience occurs. After a while, neural fatigue, and perhaps other factors, causes this state of stable equilibrium to break down, and the constant experience---say of brightness--- which has endured during the period of stable equilibrium, will abruptly cease to exist.

This process, as conceived of from a classical point of view, is purely mechanical (i.e., neuro-physiological) with consciousness the putative *consequence* of the mechanically maintained state of stable equilibrium, not the *cause* of a stable brain activity.

Once the macroscopic state of stable equilibrium that correlates with the experience comes mechanically into being it can become a controlling feature that can mechanically influence a lot of other brain activities, and one can say that the NCC exercises a sort of top-down control, which is correlated with the experience that this NCC is the correlate *of*. There could then be an "illusion" that the conscious experience *itself* is "causing" the activities that are actually being caused, purely mechanically, by the stable brain activity that is entailed by the varying NCC's.

This scenario poses the usual Hard Problem of why the perceptual experience exists at all, in view of the fact that the NCC is the sufficient physical cause of the physical action. Koch offers no solution, except his "suspicion", mentioned earlier, that the solution of the Easy Problem will produce the solution of the Hard Problem.

But, beyond that, an extrapolation of this model from this simple case to perceptions in general places a strong requirement on the neural dynamics: for every possible conscious percept there should be a potential state of stable equilibrium that maintains itself, relatively unchanging, in the presence of a range of stimuli that are sufficient to activate it.

One can of course *imagine* that a complex model of the brain exists that conforms to the principles of classical physics and that possesses all of the mechanical properties required by this classical scenario. But is it realistic to believe that such a model will someday be found, or would faithfully correspond to what is really going on, or would be useful. We know that a detailed *true* description of brain dynamics must involve the motions of ions flowing through ion channels. But then one is confronted by the fact that effects associated with Planck's constant must enter fundamentally into the dynamics, and that, in order to deal with the existence and the consequences of this universal constant of nature in a rationally coherent way, the founder's of quantum theory found themselves forced to abandon the basic precepts of classical physics, and build their theory instead upon the concept of psychophysical events. Classical mechanics emerges from that new theory only in an approximation that systematically exorcises the effects of our roles as *actors:* only in an approximation that eradicates the physical effects of the conscious choice that the founders of quantum theory found themselves forced to introduced, and that later physicists, in spite of intense efforts, have found no way to satisfactorily eliminate. In view of these developments in physics, what is the rational basis for an unwavering commitment to the claim that classical physics must be adequate.

And even if some adequate---though known to be physically false---classical model could be invented, would it be useful? A crucial recognition by the founders of quantum theory was that the essential output of a physical theory is a theoretical connection between what we can in practice choose to do and what the experienced consequences of such choices are likely to be, and that a key effect of Planck's constant is to prevent us from *ever* being able to follow, within *any* classical-type model, how such conscious inputs get transformed into experienced feedbacks. The founder's discovered that the causal gap in classical description entailed by the nonzero value of Planck's constant opened the way to a rational and useful replacement of the in-principle-inadequate classical model by a theory that explicitly brings into the dynamical structure the empirically accessible phenomenal facts concerning what we chooses to do and what we thereby learn. Why should neuroscientists not exploit such an empirically connected and potentially useful development in basic physics?

In the quantum ontology the perceptual moments are *intrinsically* discrete. The bud-like all-or-nothing character of perceptions need not be explained by imposing highly detailed, and perhaps impossible-to-meet, requirements on a classically conceived continuous neural dynamics. The discrete bud-like character of conscious events is dictated by mathematical requirements associated with the basic structure that the quantum physicists introduced in order to deal in a rationally coherent way with the limitations generated by the empirically founded need to replace numbers by operators. According to this quantum model, most of what goes on in the brain is controlled by the mechanical Process 2, which is the quantum mechanical counterpart of the single classical mechanical process. But this continuous process cannot rationally be the whole story, because it leads to physical states that are structurally different from what we perceive. To deal with this *logical* problem of the disparity between the smeared out nature of the quantum state, entailed by Planck's constant, and our relatively well defined states of knowledge about the objective world, supplied by our perceptions, orthodox quantum theory introduces discrete psychophysical events that bring the state of the physically described world into concordance with discrete increments of knowledge. Von Neumann's mathematical formulation of the quantum laws explicitly displays this discreteness, and exposes the mathematical difficulties that would ensue if one tried to replace the discrete events by a continuous process.

The physics-based model postulates, in general agreement with Koch's model, that a mind-brain event will occur when some physically described threshold is reached. But this abrupt all-or-nothing character of the event is now dictated by general mathematical requirements, not by imposing special dynamical demands upon a continuous physical process. Moreover, the empirical connections specified by the theory naturally accommodate the possibility that this psychophysical event will have a felt quality of "evaluation", perhaps a felt interest, that can trigger a rapid repeat of the discrete event. This would generate another repeat, etc., leading to a rapid sequence of essentially identical events. Such a sequence would hold in place a template for action, and would be experienced, just as we experience a movie of a static scene, as a prolonged constant experience. The constancy of the experience comes, however, from basic principles, not from some hypothesized dynamics that cannot actually be true for the same reason---the nonzero value of Planck's constant---that, according to physical principles, lies at the origin of the perceived discreteness.

<u>The wagon wheel illusion</u>

A closely connected phenomenon was discussed by Koch in the same Chapter 15. It is the "wagon wheel illusion". Every cowboy-movie goer is familiar with the fact that the wheels of a moving wagon sometimes seem to be turning backward. The explanation is simple: the flickering sequence of frames catches the wheels at times such that a later spoke has almost reached the position that an earlier spoke was at in the previous frame.

But there is a related empirical fact the needs explaining: The same illusion occurs in broad daylight!

At this point (p.264) Koch says "The implicit assumption up to this point has been that you and I experience the world in a continuous fashion: that the seamless nature of perceptual experience is reflected in the smooth waxing and waning of the NCC".

He goes on to say: "This is not the only possibility. Perception might well take place in discrete processing epochs, *perceptual moments, frames, or snapshots."…* "Within one such moment the perception of brightness, color, depth, and motion would be constant."

Koch goes on to say "Plenty of psychological data favors discrete perception, with the duration of each snapshot being quite variable, lasting anywhere between 20 and 200 msec." He cites an extreme example that was experienced by the neurologist Oliver Sacks. During a visual migraine Sacks did not see "movements as continuous but rather as a succession of "stills," a succession of different configurations and positions, but without any movement in-between, like the flickering of a film … run too slow."

Koch's explanation of the evidence in favor of discrete perception, in terms of what is going on in the continuously changing brain, harks back to his explanation of the constancy of the perception of the steady light. He says (p. 264): "Activity at the essential mode for some attribute would build up until a dominant coalition established itself and the NCC came into being. If the subject continued to attend to the stimulus, the dynamics of the system would have to be such that with some degree of regularity the NCC turn off and on again, constant within one perceptual moment but changing from one to the next before reaching a new quasi steady-state."

Again, Koch's explanation of this fundamentally bud-like character of perception depends upon very special vaguely imagined putative properties of the mind-brain, conceived of in classical mechanistic materialist terms that are known to be fundamentally false. On the other hand, the explanation of this discreteness within the framework of the empirically validated principles of quantum physics stems not from conjectured special properties of the classical approximation, which eliminates all the physical effects of our thoughts, but rather from mathematical demands arising from the need to incorporate rationally into our understanding of all natural phenomena, and most particularly ones depending significantly on atomic-level dynamics, the far-reaching effects of the universal constant of nature discovered at the end of the nineteenth century by Max Planck.

References

Bohr, N. (1934). Atomic Theory and the Description of Nature. Cambridge: Cambridge University Press.

Bohr, N. (1958). Atomic Physics and Human Knowledge. New York: Wiley.

James, W. (1911). "Some Problems in Philosophy" Chapter X

Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. (Englewood, CO: Roberts and Co,)

Schwartz, J, Stapp, H, & Beauregard M. (2005). "Quantum Physics in Neuroscience and Psychology: A Neurophysical Model of Mind/Brain Interaction," *Philosophical Transactions of the Royal Society* B 360 (1458) 1309-27.

Stapp, H. (2004). *Mind, Matter, and Quantum Mechanics* (Berlin, Heidelberg, New York: Springer) Second edition. Chapters 6 & 12.

Stapp, H. (2005). "Quantum Interactive Dualism: An Alternative to Materialism," J. Conscious Studies.12, No.11 pp. 43-58.