Is Quantum Mechanics Relevant To Understanding Consciousness?

A Review of *Shadows of the Mind* by Roger Penrose

Stanley A. Klein
Department of Vision Sciences
University of California
Berkeley, CA 94720
USA

klein@adage.berkeley.edu

PSYCHE, 2(3), April 1995
http://psyche.cs.monash.edu.au/v2/psyche-2-03-klein.html

KEYWORDS: consciousness, duality, Libet, metaphysics, Penrose, quantum mechanics, reductionism.

REVIEW OF: Roger Penrose (1994) *Shadows of the Mind*. New York: Oxford University Press. 457 pp. Price: $US 25 hbk. ISBN 0-19-853978- 9.

1. Introduction

1.1 The present essay explores three issues raised by Penrose in *Shadows of the Mind* (abbreviated *Shadows* from here on): (1) is classical (non-quantum) science incapable of understanding brain operation?; (2) are long-range quantum effects able to produce measurable changes in neural activity?; (3) why have so many researchers proposed a strong connection between quantum mechanics and consciousness? In connection with this third topic, I will argue that although Penrose is probably wrong about the physics of quantum mechanics being relevant to the (third person) neural correlates of awareness, the metaphysics of quantum mechanics may be essential to understanding the (first person) subjective nature of consciousness. In Penrose's approach these two aspects become inseparably intertwined, adding confusion to an already murky area.

2. The Presumed Deficiencies Of Classical Mechanics

2.1 The first half (Part I) of *Shadows*, titled "Why We Need New Physics to Understand the Mind" clarifies Penrose's belief that classical, algorithmic neural dynamics is insufficient for understanding the operation of the brain. I will not examine his main argument, based on Gödel's theorem, here, as it has been widely discussed elsewhere. Instead I will examine several non-Gödel arguments regarding the purported inadequacy of classical mechanics for understanding brain operation.

2.2 In his review of *Shadows*, Wilczek (1994) gave Penrose an important challenge. Rather than searching for a Gödelian feat that humans, but not robots, can do, Wilczek suggests merely looking for perceptual feats that humans can do more efficiently than robots. As Penrose points out in Section 7.3, quantum computing can be much faster than classical computing. Perception is a good area to examine since evolution has been perfecting perception for quite a while, making it an area in which humans should excel. In my own area of research, visual perception, I am often confronted with counterintuitive illusions and surprisingly rapid visual effects. However, I have always been able to account for the data using standard, classical neural models. I have never come across data from any sort of mental processing (other than non-verifiable ESP effects) that could not, in priciple, be accounted for by clever, non-quantum, physiologically plausible neural models. While it is true that at the present time robots do poorly at high level classifications such as face recognition, few researchers would say that future classical robots will be incapable of doing such tasks well and rapidly.

2.3 At the end of Chapter 7, Penrose cites two studies of Libet and colleagues (see Libet, 1993 for an excellent summary) as providing evidence for the non-classical nature of consciousness. The studies were concerned with the timing relationships between brain activity, conscious awareness, and physical action (sensory stimulation in one set of experiments and motor response in a different set of experiments). Of these experiments, Penrose says:

> we seem to be driven to the conclusion that in any action in which an external stimulus leads to a consciously controlled response, a time delay of some one and one-half seconds would seem to be needed before that response can occur. For awareness would not even take place until half a second has passed; and if that awareness is to be put to use, then the apparently sluggish machinery of free will would then have to be brought into play, with perhaps another second's delay.

2.4 Penrose believes (correctly) that a 1.5 second delay like this would contradict our experience. He interprets this contradiction as providing evidence for quantum processing (p. 388):

> ...if, in some manifestation of consciousness, classical reasoning about the temporal ordering of events leads us to a contradictory conclusion, then this is a strong indication that quantum actions are indeed at work!

2.5 As far as I can tell, however, nothing in Libet's experiments comes close to requiring quantum explanations. Before invoking exotic explanations it is useful to reconsider Libet's two experiments to see whether Penrose's estimate of a 1.5 sec response time is valid. (The following analysis has been developed in collaboration with my colleague Marcus Baldo.)

2.6 Consider first Libet's experiments with skin and cortical stimulation that Penrose takes as evidence for a 500 msec delay. Libet (1993) is clear that this long delay is only appropriate for stimuli that are near threshold. He points out (p. 131) that a cortical stimulus of about 100 msec would likely elicit awareness. The time of the actual awareness event is very difficult to assess (as was pointed out by Ian Glynn (1990) in his very useful commentary on Penrose's (1989) earlier treatment of Libet), but psychophysical experiments such as that by Nijhawan (1994) lead me to believe that the time delay between the sufficient stimulus and conscious awareness is less than 100 msec. I know of no evidence that would imply the delay between a suprathreshold stimulus and conscious awareness is greater than 100 msec.

2.7 Consider now the evidence that there is a 1-second delay from conscious awareness to motor output. Penrose cites Libet's experiments with volitional finger flexing, in which the moment of decision was determined from a subject's report of the position of a rapidly rotating clock hand. According to Libet (1993, p. 128), however, the subjective awareness of the decision to act occurred only 200 msec before the motor act with substantial unconscious processing taking place before that. No surprises there. Furthermore, Libet's use of the rotating clock can overestimate the timing of awareness, as Nijhawan's (1994) experiments, also using a rotating clock, showed. In any case, Libet's experiments are not relevant to Penrose's calculation of a 1.5 sec delay: the latter is related to a stimulus-response situation, whereas Libet's experiments involved volitional motions.

2.8 It would be useful for Penrose, in his response, to clarify his calculation of the 1.5 sec delay, since a 300-400 msec delay seems fully compatible with Libet's data.

2.9 One last comment on Penrose's discussion of the presumed deficiencies of classical mechanics. On pages 372-373 of *Shadows*, Penrose discusses the global nature of consciousness:

> The unity of a single mind can arise, in such a description, only if there is some form of quantum coherence extending across at least an appreciable part of the entire brain.

2.10 In Penrose's approach the unified feeling of consciousness is associated with long range quantum coherence, but why can't the required coherence be achieved classically? Classical neural networks with feedback can produce surprisingly rich, coherent activity. With signals going from one neuron to the next in a couple of milliseconds (and even faster dendritic processing) it is possible for coherent, classical activity to spread across the brain on a 10 msec time scale.

3. Penrose's New Physics, Microtubules, And Mind

3.1 Even though we are not driven to quantum mechanics because of an inability of classical mechanics to deal with neural activity, it is still worth exploring Penrose's quantum ideas. For one thing, there is the intriguing link between quantum mechanics and the role of the observer. I will deal with Penrose's quantum ideas in two parts. This section will examine Penrose's ideas on the connection of quantum physics and brain operation. The next section examines quantum metaphysics and subjective awareness.

3.2 For a brief, elegant introduction to the laws of quantum mechanics one cannot beat Feynman (1985).

There are two steps in quantum calculations. First, one calculates the amplitude for the outcome of an experiment. This calculation is achieved using the Feynman rules governing the interaction of "particles" such as electron and quarks with "forces" such as photons and gluons (Feynman, 1985). These rules are similar to rules governing waves. Second, the square of the wave amplitude gives the probability of finding a particular outcome of particle locations. A major question is: at what point does the wave amplitude get reduced (Penrose's R process) to particle probabilities? That is, when does one apply the second step? Penrose devotes Chapter 6 to a discussion of whether the R process is real and to several alternative interpretations of R. Penrose believes that the reduction of the wave function is governed by a quantum gravity effect that occurs at intermediate scales of size between microscopic atoms and macroscopic brains. I have no hesitation in recommending Chapter 6 for anyone interested in gaining a deep understanding of the quantum mysteries and the possibility of novel solutions. My problem comes when he attempts to apply these ideas to the brain in Chapter 7.

3.3 In order to clarify the problems facing Penrose, an example will help. Suppose a neuron is in a quantum superposition of states A (the neuron fires) and B (the neuron doesn't fire). The superposition of states can be written as: a|A> + b|B>, where |a|^2 and |b|^2 are the probabilities that the neuron fires or doesn't fire. For convenience we are using Dirac's notation, |A> to represent the state A. A critical difference between classical mechanics and quantum mechanics is that in the latter the relative sign (or phase) of the amplitudes a and b can lead to measurably different outcomes if the overlap between states |A> and |B> becomes substantial as they evolve in time. The measurable difference is referred to as quantum interference. In classical mechanics the relative sign (phase) of a and b is not measurable since only the probabilities |a|^2 and |b|^2 ever enter classical calculations. Thus the question of whether there are measurable quantum effects that differ from classical predictions is directly connected with the question of overlap of states. For our example with a neuron's firing I suspect everyone will agree that the biochemical difference produced by firing is too large to ever allow overlap between the fired and unfired state. Thus measurable quantum effects would be expected to take place before a neuron fires.

3.4 It will be difficult to find quantum effects in pre-firing neural activity. The big problem facing quantum superpositions in the brain is that the brain operates at high temperatures (there are constant thermal vibrations of the proteins and other molecules involved with neural activity) and is made of floppy material (the neural proteins can undergo an enormously large number of different types of vibration). Furthermore, it is highly likely that states |A> and |B> will get entangled with different modes of oscillation of the environment. When that happens it is exceedingly unlikely that states |A> and |B> will ever develop substantial overlap, so the intrinsically quantum effects (the deviations from classical predictions) become negligible. The challenge facing Penrose is to convince the reader that there is much less "floppiness" than appears to be the case.

3.5 To meet this objection, Penrose makes wishful conjectures about properties of microtubules. These proteins, found in all cells, have properties that could conceivably be useful for computation within individual neurons (Hameroff & Watt, 1982). But while it may be conceivable that microtubules maintain quantum coherence within a neuron, it is difficult to imagine how this coherence can be maintained across neurons. Penrose is aware of this problem when he says (p. 372): "... the quantum coherence must leap the synaptic barrier between neuron and neuron. It is not much of a globality if it involves only individual cells!"

3.6 What Penrose's story lacks is an account of how quantum coherence can leap the synaptic barrier. All he says is that evolution is clever, and maybe it has found a way to achieve this impossible sounding feat. He discusses how microtubules can alter synaptic strengths in an interesting way, but nowhere is there any discussion of the nature of synaptic modulations that can be achieved quantum-mechanically but not classically. The quantum nature of neural activity across the brain must be severely restricted, since Penrose concedes that neural firing is occurring classically. I kept rereading Sections 7.6 (Microtubules and consciousness) and 7.7 (A model for a mind) in hopes of finding a plausible argument for how a coherent quantum state could be preserved across the brain. I thought that Penrose might invoke the .5 cm microwaves that have been associated with microtubules, but he was too smart for that. He must have felt that invoking microwaves to achieve quantum coherence across neurons would be too easily disproved. For his hypothesis to have any chance of success, Penrose needs new mechanisms of neural communications other than the presently known electrochemical mechanisms, but given the explanatory power of presently-known mechanisms, it is unlikely that neuroscientists will mount a search for these new and exotic mechanisms soon.

4. Physics vs. Metaphysics

4.1 Given the skepticism I expressed in the preceding section about Penrose's Chapter 7 "Quantum Theory and the Brain', it might seem that I am skeptical about all aspects of Penrose's attempt to connect quantum mechanics and consciousness. Not so. I believe that the metaphysical underpinnings of quantum mechanics are crucial for a deep understanding of consciousness and the connection between mind and brain. The distinction being made is between science (asking testable questions) and metaphysics (asking nontestable questions). The metaphysics of quantum mechanics, with its insight into the role of the observer, may well be relevant to the subjective aspects of how studies of mind can fit into the scientific worldview.

4.2 Chapters 5 and 6 of *Shadows* provide a wonderful presentation of the paradoxical nature of quantum mechanics. The framework that Penrose develops in those chapters has direct applications to the metaphysics of the mind-brain connection. The challenge is to find a satisfactory way to associate the "observer" of subjective awareness with the "observer" of quantum mechanics (Penrose's R process). Stapp (1993) and Penrose would like this association to be in the realm of science; Klein (1992, 1993) has argued that it could be in the realm of metaphysics.

4.3 The big problem in the metaphysics of quantum mechanics is the question of where to place the split between the observer and the observed. The astonishing finding of von Neumann (1955) is that its placement is irrelevant to any measured event. The Feynman rules for the world below the split and the classical rules for the world above the split are so clever that the split is moveable. This is the brilliant manner in which the quantum duality avoids the difficulties encountered by the previous dualities of Plato, Descartes and Kant. Previous dualities contained inconsistencies when the two sides were compared. There are no inconsistencies between the two halves of the quantum duality. Present quantum theory, with its flexible split placement, allows the neural correlates of awareness to be above the split (the neural correlates of awareness become the observer) and the remaining (unconscious) neural activity to be below. This is the placement advocated by von Neumann (1955), Wigner (1961) and Stapp (1993). Stapp, in particular, has been lucid in writing about the conscious act being connected with the reduction process.

4.4 In Penrose's response to the present articles it would be helpful to see a clarification about which aspects of the quantum-consciousness connection he believes are within science and which are within metaphysics. Even if it finally turns out that only the metaphysical aspects of quantum mechanics are relevant to consciousness, *Shadows of the Mind* would still be an excellent resource for gaining an appreciation for those aspects.

Acknowledgements

I would like to thank Marcus Baldo, Henry Stapp, David Chalmers and Dan Rokhsar for useful comments.

References

Feynman, R.P. (1985). *QED.* Princeton: Princeton University Press.

Glynn, I.M. (1990). Consciousness and time. *Nature, 348,* 477-479.

Hameroff, S.R., & Watt, R.C. (1982). Information processing in microtubules. *Journal of Theoretical Biology, 98,* 549-561.

Klein, S.A. (1992). Duality of psycho-physics. In A. Gorea, ed., *TacIt Assumptions in Vision Research.* Cambridge: Cambridge University Press.

Klein, S.A. (1993). "Will Robots See?" In *Spatial Vision in Humans and Robots.* Harris and Jenkin eds. Cambridge: Cambridge University Press.

Libet, B. (1993). The neural time factor in conscious and unconscious events. *Experimental and Theoretical Studies of Consciousness* (CIBA Foundation Symposium 174) 123-146. Chichester: Wiley.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370,* 256-7

Penrose, R. (1989).*The Emperor's New Mind.* Oxford University Press.

Stapp, H.P. (1993). *Mind, Matter and Quantum Mechanics.* Berlin: Springer-Verlag.

von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics.* Princeton: Princeton University Press.

Wigner, E.P. (1961). Remarks on the mind-body question. Reprinted in Wheeler and Zurek, eds., *Quantum Theory and Measurement.* Princeton: Princeton University Press.

Wilczek, F. (1994). A call for a new physics.*Science, 266,* 1737 - 1738.

Return to PSYCHE home page