©1992 The Institute of Mind and Behavior, Inc.
The Journal of Mind and Behavior
Winter 1992, Volume 13, Number 1
Pages 29–36
ISSN 0271–0137

Quantum Theory and Consciousness

Ben Goertzel

University of Nevada, Las Vegas

This article seeks to clarify the relation between consciousness and quantum physics. It is argued that, in order to be consistent with quantum theory, one must never assert that conscious action has caused a given event to occur. Rather, consciousness must be identified with "measurement" or, more concretely, with an increase in the entropy of the probability distribution of possible events. It is suggested that the feeling of self-awareness may be associated with the exchange of entropy between groups of quantum systems which are so tightly coupled as to be, for all practical purposes, an indivisible unit. Such groups of systems may be understood to measure themselves. Two interpretations of the quantum theory of consciousness are distinguished: one in which consciousness is defined as quantum measurement; and one in which this measurement is hypothesized to correlate with a certain biological phenomenon called consciousness.

For sixty years physicists have struggled with the paradox of quantum measurement. However, despite a number of theoretical advances, rather little progress has been made toward resolution of the basic dilemma. According to quantum physics, no physical entity is ever in a definite state; the most one can ever say about a given entity is that it has certain probabilities of being in certain states. And yet, both in daily life and in the laboratory, things do appear to have definite states.

For instance, the equations of quantum physics predict that, in certain situations, an electron has a 50% "chance" of having a positive spin, and a 50% "chance" of having a negative spin. Yet when the physicist probes the electron in the laboratory, it appears to have either a positive spin or a negative spin. According to the equations of quantum physics — the Heisenberg equation or the Schrödinger equation — such a reduction to a definite state is impossible.

Of course, one may have various degrees of probabilistic certainty. In certain situations, one might know an electron to have a 90% "chance" of hav-

---

Requests for reprints should be sent to Ben Goertzel, Ph.D., Mathematics Department, University of Nevada, Las Vegas, 4505 S. Maryland Parkway, Las Vegas, Nevada 89154.

ing positive spin, and a 10% "chance" of having negative spin. But there can never be 100% definite knowledge. Heisenberg's indeterminacy principle says that one can never have complete knowledge of the state of any particle: the greater the accuracy with which one knows its position, the less the accuracy with which one can know its momentum; and vice versa. In order to predict what the particle will do in the future, one needs to know both the position and the momentum; but according to quantum physics, this is possible only probabilistically.

This sort of indeterminacy is a proven scientific fact, in that quantum theory is the only known theory that correctly explains the behavior of microscopic particles, and it predicts only probabilities. Classical mechanics and electromagnetism gave definite answers about the behavior of microscopic particles, but these answers were experimentally wrong. And it seems likely that, if quantum theory is someday superseded, the theory which follows it will build on the probabilistic nature of quantum theory, rather than regressing to classical ideas. In fact, it has been proved mathematically (Bell, 1987) that any physical theory satisfying certain simple requirements must necessarily have properties similar to those of quantum theory: it must deal only in probabilities.

In his classic analysis of quantum theory, John von Neumann (1955) introduced the "projection postulate," an addition to the basic principles of quantum physics which states simply that, when an entity is measured, it reduces to a definite state. This approach appears to be adequate for most practical problems of quantum mechanics; and, although, many physicists find it unacceptable, there is no equally elegant alternative. The trouble is that no one has ever given a fully satisfactory characterization of the "measurement" process (Omnes, 1990).

Originally it was thought that a microscopic event could be considered to be measured when it "registered" an effect on some macroscopic entity. The justification for this was the belief that, as entities become larger and larger, the probabilistic nature of quantum physics becomes less and less relevant to their behavior. For instance, according to quantum physics a baseball dropped from a window has an infinity of possible paths, but it is intuitively obvious that one of them, or one small class of them, is more likely than the others.

But this naive identification of measurement with macroscopic effect cannot stand up to criticism. Spiller and Clark (1986) have constructed a Superconducting Quantum Interference Device (SQUID) which is about the size of a thumbnail and yet displays the same sort of uncertainty as a spinning electron. One can never know both the intensity and the flux of its magnetic field with perfect accuracy; there is a finite limit beyond which further accuracy is impossible. Its state is fundamentally a probabilistic superposition.

And it appears that the brain may display a similar form of quantum indeterminacy (Changeaux, 1985). Recall that a neuron fires when its charge exceeds a certain threshold amount. It follows that, on occasion, highly unpredictable quantum phenomena may push the charge of a neuron over the threshold. And this neuron may then set other neurons off, and so on — in this manner a tiny quantum indeterminacy may give rise to a huge neurophysiological uncertainty. If the extra charge has a fifty-fifty chance of being there, then the entire pattern of neuronal firing that ensues from its presence has a fifty-fifty chance of being there. This pattern of neuronal firing might, for instance, represent a state of mind. And when you consider the fact that there are over a hundred billion neurons in the brain, the possibilities for interlocking quantum uncertainties are astounding. The exact numbers are difficult to estimate, but it appears that this may be a significant phenomenon.

One intriguing alternative to the projection postulate is Everett's (1957) "many-worlds hypothesis," which assigns to each uncertain situation an array of universes, one corresponding to each possible outcome. For instance, according to the many-worlds hypothesis, every time a physicist observes an electron to have positive spin, there is an alternate universe which is exactly identical to this one, except that in the alternate universe the physicist observes the electron to have negative spin. This is an interesting possibility, but it is empirically indistinguishable from the projection postulate, since these alternate universes can never be observed.

Another alternative, first proposed by Wigner (1962), is that "measurement" may be defined as "registration into consciousness." To see the motivation for this radical idea, let us consider the infamous paradox of Schrödinger's cat (1948). Here the peculiarity of quantum theory is elevated to the level of absurdity. Put a cat in a soundproofed cage with a radioactive atom, a Geiger counter and a vial of poison gas. Suppose that the atom has a fifty-fifty chance of decaying within the hour. According to the dynamical equations quantum physics, this is *all one can know* about the atom: that it has a fifty-fifty chance of decaying. There is no possible way of gaining more definite information. Assume that, if the atom decays, the Geiger counter will tick; and if the Geiger counter ticks, the poison vial will be broken. This set-up is bizarre but not implausible; a clever engineer could arrange it or something similar.

What is the state of the cat after the hour is up? According to quantum theory without the projection postulate, it is neither definitely alive nor definitely dead but half and half. Because the atom *never* either definitely decays or definitely doesn't decay: quantum physics deals only in probabilities. And if the atom never either definitely decays or definitely doesn't decay, then the cat never definitely dies or definitely doesn't die.

One might argue that the cat is not in a state of superposition between life and death, but rather has a fifty percent chance of being alive and a fifty percent chance of being dead. But according to quantum theory without the projection postulate, the cat will never collapse into a definite state of being either alive or dead. What sense does it make to suggest that the cat has a fifty percent chance of entering into a state which it will never enter into? The function of the projection postulate is to change the statement that the cat is half dead and half alive into a statement about the probabilities of certain definite outcomes.

Of course, the fact is that if we look in the box after the hour is up, we either see a dead cat or a live cat. Somehow, by the time the observation is made, one of the two possibilities is selected — *definitely* selected. But *when*, exactly, does this selection occur? Since measurement cannot be defined as macroscopic registration, this is a very serious problem. And the problem is resolved very neatly by the hypothesis that *probabilistic occurrences are replaced by definite occurrences when they enter consciousness*.

For instance, this implies that Schrödinger's cat is not half dead and half alive, but rather has a fifty percent chance of being dead and a fifty percent chance of being alive. The cat becomes definitely dead or definitely alive *when a conscious being sees it.* As Goswami put it,

> it is our consciousness whose observations of the cat resolves its dead-or-alive dichotomy. Coherent superpositions, the multifaceted quantum waves, exist in the transcendent order until consciousness brings them to the world of appearance with the act of observation. And, in the process, consciousness chooses one facet out of two, or many, that are permitted by the mathematics of quantum mechanics, the Schrödinger equation; it is a limited choice, to be sure, subject to the overall probability constraint of quantum mathematics (i.e. consciousness is lawful) . . . . [C]onsciousness . . . *is not about doing something to objects via observing, but consists of choosing among the alternative possibilities that the wave function presents and recognizing the result of choice.* (1990, p. 142)

That is, the mind does not create the world in the sense of reaching out and physically modifying events. But it creates the world by selecting from among the wide yet limited variety of options presented to it by the probabilistic equations of physics.

Consciousness and Nonlocality

The measurement paradox is only one of the philosophically troublesome aspects of quantum physics. Bell's Theorem (1987), with its implication of instantaneous communication between distant events, is equally unsettling. The simplest example is the Einstein–Podolsky–Rosen thought experiment, in which two electrons,initially coupled, are shot off in different directions. It is assumed that each one flies for millions of miles without hitting anything. Each one, according to quantum physics, has a fifty-fifty chance of spinning to the right or to the left — there is no way to make a more accurate prediction. However, the rules of quantum physics *do* imply that the two are spinning in opposite directions: if one is spinning to the right, then the other one is spinning to the left; and vice versa.

Now suppose someone measures *one* of the electrons, so that it all of a sudden assumes a definite value. Then the other electron will *immediately* also assume a definite value because it is known that the two are spinning in opposite directions. If one is measured to be spinning to the right, then the other is instantaneously known to be spinning to the left. When Einstein conceived this example, he thought he had disproved quantum mechanics because nothing so absurd could possibly be true. After all, he asked, *how does the one electron tell the other one which way to spin?* Special relativity forbids information from traveling faster than the speed of light.

But, absurd as this may be, it is an experimentally proven fact (Aspect, Grangier, and Roger, 1982). Scenarios very similar to the original Einstein–Podolsy–Rosen thought experiment have been tested in the lab. Technically speaking, this peculiar "nonlocality" does not contradict special relativity, because no information is being transmitted, only a *correlation*. But this hardly makes it any easier to comprehend.

Reality does not consist of pairs of electrons, coupled and then shot out into space a million miles in opposite directions. Consider the following thought experiment. Split apart two coupled physical systems, say A and B. Suppose that, from the state of A, one could infer the state of B, and vice versa. Leave A alone but let B interact with C for a while, and then separate B from C. Finally, measure A. A is collapsed into some definite state. If B had not interacted with C, one could say that the state of B would also, immediately, collapse into some definite state. But the state of B now depends also upon the state of C, which according to quantum physics has no definite value but is rather an array of possibilities. So the measurement of A does not collapse B to a definite state. But it does, however, decrease the uncertainty involved in the state of B. It increases the "closeness" of B to a definite state.

Technically speaking, assume that $p=(p_1,p_2,\ldots,p_n)$ denotes the probabilities of the various possible states in which B might be. Then one may show that, in the situation described above, the measurement of A necessarily changes p into a new set of probabilities $p'=(p_1',\ldots,p_n')$ so that $H(p_1,\ldots,p_n) < H(p_1',\ldots,p_n')$, where H is the entropy function

$$H(p_1,\ldots,p_n) = -[p_1\log p_1 + \ldots + p_n\log p_n]$$

A similar statement may be made when the possible states of B form a continuum rather than a discrete set. Recall that the entropy of a probability distribution is a measure of its uncertainty, or its distance from the most certain distribution. This thought experiment may be generalized. What if the state of B *cannot* be completely determined from the state of A? If the state of A yields any information at all about the state of B, then it is plain that the same result holds. If A and B were ever coupled, no matter how loosely, no matter what they have done since, measurement of A reduces the uncertainty of the probability distribution characterizing the states of B. Bell's Theorem implies that this sort of propagation of certainty is a necessary aspect of any physical theory that is mathematically similar to quantum theory.

In terms of consciousness, what does this mean? A little consciousness can go a long way! If two sets of possibilities have been coupled in the past, and are then separated, whenever consciousness makes one of them definite, the other one becomes definite automatically, instantaneously, without any physical causation involved.

Quantum Consciousness: Philosophy or Science?

The introduction of consciousness provides a philosophically elegant resolution of the paradox of quantum measurement. But in a way it is an abuse of the word "consciousness." What qualities does this abstract entropy-decreasing consciousness share with our common-sense understanding of consciousness?

For instance, Mandler has proposed that

> . . . [C]onscious constructions represent the most general interpretation that is appropriate to the current scene in keeping with both the intentions of the individual and the demands of the environment . . . . Thus, we are aware of looking at a landscape when viewing the land from a mountaintop, but we become aware of a particular road when asked how we might get down or of an approaching storm when some dark clouds "demand" inclusion in the current construction. In a problem-solving task, we are conscious of those current mental products that are closest to the task at hand, i.e. the likely solution to the problem. (1985, p. 81)

Whether or not this particular formulation is exactly correct, it seems plain that some similar characterization must hold true. Consciousness seems to have a role in planning and decision-making, but it is rarely involved in the minute details of everyday life: walking, turning the pages of a book, choosing words in conversation, doing arithmetic with small numbers, etc.

The decision-making aspect of consciousness is intuitively harmonious with quantum theory: in making a decision, one is reducing an array of possibilities to one definite state. There is a sense in which making a decision corresponds to selecting one of many possible universes. But the quantum theory of consciousness gives us no indication of why certain decisions are submitted to consciousness, but others are not.

One of the main problems here is that it is not clear what function the quantum theory of consciousness is supposed to serve. In its standard form, as presented by Wigner (1962) or Goswami (1990), consciousness is *defined* as the reduction to a definite state, or more generally as the decrease of the entropy of an array of possible states. This interpretation provides a transcendentalist resolution of the mind-body problem, made explicit by Goswami when he suggests that, as a heuristic tool, we consider the mind to be a coupling of two computers, a classical computer and a quantum computer. The quantum computer behaves in a way which transcends ordinary biophysics, and it is this transcendence which is responsible for consciousness.

But there is another, more radical, way of interpreting the quantum theory of consciousness. One may begin with the assertion that consciousness is a *process* which is part of the dynamics of certain physical systems, e.g., human brains. This means that consciousness has some direct physical effect: that, for instance, when a pattern of neural firings enters consciousness, consciousness changes it in a certain characteristic way. This is the way the neuroscientist thinks of consciousness (Changeaux, 1985).

Given this biological characterization of consciousness, one may then hypothesize that the quantum-theoretic entropy reduction of arrays of possible states is *correlated with* the biological process of consciousness. This point of view places less responsibility on quantum theory than does the interpretation of Wigner and Goswami: it does not require quantum theory to explain psychological facts. Rather, it portrays consciousness as the point of connection between psycho-biological dynamics and physical dynamics; the bridge between the mind and the world. This provides an approach to the mind-body problem which is rather different from Goswami's transcendentalism.

Self-Awareness

In conclusion, I would like to point out that the quantum view of consciousness yields an interesting interpretation of that intangible feeling of self-awareness that accompanies consciousness of external objects or definite ideas. Consider the following scenario. P and Q are closely coupled algorithms, each one continually modifying the other. Simultaneously, consciousness greatly reduces the uncertainty of both the distribution of possible states of P and the distribution of possible states of Q. The reduction of the uncertainty P then reduces the uncertainty of Q yet further; and vice versa. The result is that the combined entity $P \cup Q$ has, in effect, *looked at itself* and reduced its own entropy.

It is not justifiable to say that $P \cup Q$ did not really look at itself, that what really happened was that P and Q looked at each other. Because according to quantum physics, if we *observed* $P \cup Q$ to see what was really happening, this would change the probability distributions. P and Q are quantum coupled, and this means they are effectively one entity. Clearly, this situation is not rare: feedback between different prominent structures is probably not the exception but the rule.

According to this analysis, the feeling of self-awareness is not logically inherent to consciousness; it is rather an extremely common by-product of consciousness. This accounts for the fact that we are not continually absorbed with the sensation of self-awareness: it flits in and out of consciousness. Self-awareness is not quite the same as consciousness, but the two are inextricably interlinked.

References

Aspect, A., Grangier, D., and Roger, G. (1982). Tests of Bell's inequality with pairs of low-energy correlated photons. *Physical Review Letters*, *49*, 91–97.

Bell, J.S. (1987). *Speakable and unspeakable in quantum mechanics*. Cambridge: Cambridge University Press.

Changeaux, J.P. (1985). *Neuronal man*. New York: Oxford.

Everett, H. (1957). "Relative state" formulation of quantum mechanics. *Review of Modern Physics*, *29*, 454–462.

Goswami, A. (1990). Consciousness in quantum physics and the mind-body problem. *Journal of Mind and Behavior*, *11*, 75–97

Mandler, G. (1985). *Cognitive psychology: An essay in cognitive science*. Hillsdale, New Jersey: Erlbaum.

Omnes, R. (1990). Some progress in measurement theory: The logical interpretation of quantum mechanics. In W.H. Zurek (Ed.), *Complexity, entropy and the physics of information* (pp. 495–512). New York: Addison–Wesley.

Schrödinger, E. (1948). The present situation in quantum mechanics. *Proceedings of the American Philosophical Society*, *124*, 323–328.

Spiller, T., and Clark, T. (1986). SQUIDs: Macroscopic quantum objects. *New Scientist*, December 4, 36–40.

von Neumann, J. (1955). *Mathematical foundations of quantum mechanics*. Princeton: Princeton University Press.

Wigner, E.P. (1962). *Symmetries and reflections*. Bloomington: Indiana University Press.