Appendix F. Science, Free Will, and Morality.

The topic of this appendix is the core topic of the entire book. But it may be useful at this point to distill out the essence of the argument presented in the main text, and to expand upon it, by contrasting it with the principal contemporary alternative approach to the subject.

Advances in the scientific understanding of the connection between our minds and our brains can be expected to have important applications in the treatment of medical and psychological aliments. But an issue of far greater ultimate significance is the impact of the new developments in science on the foundations of moral philosophy. For at the present stage of technological development it is human values that largely control the destiny of our species.

For the past few centuries the main impact of science on moral issues has stemmed from the fact that the concepts of classical physics entail a materialist conception of nature in which each human being is reduced to a collection of cogs in a giant machine. Each of us becomes, according to that view, a puppet whose every act is controlled by the motions of mindless microscopic entities, which themselves are ruled by mechanistic laws. By *mechanistic* I specifically mean “local deterministic”: each elementary entity, at each instant of time, is located essentially at a point in space, and its immediate future is completely fixed by the physical states of its immediate neighbors. Thus the entire course of world history, from big bang to big crunch (or forever, if the universe lives forever), is fixed already at some early primordial time. Such a mechanistic universe is, according the way most people use the words “free will”, quite devoid of that property. Thus just as we normally say that a mechanical robot lacks free will, so must we say the same of ourselves, insofar as that mechanistic/materialistic world view is accepted. But how can we reconcile this “conclusion of science” with the notion of personal responsibility that normally underlies our moral precepts. That formidable problem is immediately raised by an adoption of the world view of classical physics.

Certain philosophers have resolutely set themselves the task of reconciling the mechanistic materialism of classical physics with the concepts of free will and personal responsibility that underlie our

usual precepts of morality. No one has tried harder than Daniel Dennett. In his recent book *Freedom Evolves* he views much of his 30-year output of books, from his 1984 book *Elbow Room*, through *Consciousness Explained* and *Darwin's Dangerous Idea,* as the construction of the foundations for his attack on this difficult problem. But, hard as he has tried, people have seemed incapable of properly understanding his main points, which are that "Our minds are just what our brains non-miraculously do,…"(p.xi) and that this premise is compatible with conscious "free will".

Dennett's main premise is a simple one, namely the "identity theory" thesis that each conscious thought is the very same thing as some brain activity, and, moreover, that brain activities can be regarded as being---insofar as they bear on these issues---governed by the mechanistic-deterministic laws of nineteenth century classical physics. But a clear understanding of Dennett's approach seem to be eluding even his most sympathetic readers. He is frustrated (p. 20) with the defection of Steven Pinker, whom he had classified as one of the "responsible, cautious naturalists" like himself, but who, Dennett now says, continues to dally with what he describes as "mysterian doctrines of consciousness". Dennett is disappointed also with the fact that Robert Wright, who he says gives a fine presentation of most of the ideas that he will be presenting, finds himself unable to fully support Dennett's "uncompromising materialism". "Mysterian doctrines of consciousness" appear to be the notion that our conscious thoughts and feelings---those elements of our the streams of consciousness that are the only realities that we actually *know*---are non-identical to the theoretical inventions of that great alchemist, Isaac Newton; and "uncompromising materialism" appears to mean equating any deviation from nineteenth century classical mechanistic determinism to "defying the laws of physics" (p.1), in spite of the contrary findings of twentieth-century physicists.

Dennett mentions also a book by Richard Dooling that includes an "insightful and accurate" précis of his (Dennett's) theory of conscious, but then "gets the part about free will dead wrong, *just the way that some neuroscientists have done*." (Dennett's italics.) Thus Dennett, in spite of his intensive long-term effort to explain and defend his ideas, seem to be unable to get even some very serious and like-minded people to correctly understand or agree with him.

So what is going on here? The people Dennett cites seem to be sympathetic with his ideas, but then, in the end, cannot understand him, or disagree with him, or even get key points dead wrong. Why are his ideas so hard to communicate to others?

Dennett (p. 224) quotes Tom Wolfe as noting that he, Dennett, (along with E. O. Wilson and Richard Dawkings) are "presenting elegant arguments" as to why mechanistic materialism does not diminish our self-image, but that the message "is not rippling out to the public. …The conclusion people out beyond the laboratory are drawing is: *The fix is in! We are all hardwired!* That, and: *Don't blame me! I'm wired wrong!"*

Dennett notes (p. 226) that he has already discussed this matter of the status of free will and morality in *Consciousness Explained*, "but that discussion was obscure and difficult and needs refreshing." However, the extreme difficulty with Dennett's position, testified to by the fact that he has written books and books about it yet smart and sympathetic readers still don't get it, suggests that maybe he is trying, with great cleverness indeed, to prove true a false idea, namely the compatibility of (1), the idea that each of us is a mechanically controlled automaton, with (2), the claim that each of us has a free will that adequately undergirds personal responsibility and our ability to act morally. Dennett says (p. 223) that "I've finally come to the conclusion that some people like the confusion." But the people involved here include intelligent philosophers and scientist who are diligently striving to root out confusion. Why are thirty years of books insufficient to get his thesis across to people such as these?

Dennett's view stems from a basic commitment to the mechanistic determinism of classical physics, coupled with the idea that consciousness is not an idle bystander. These commitments lead naturally to "identity theory", to the idea that consciousness is not a *partner* with matter, but *is, itself,* mechanistically deterministic brain-matter in action. His efforts to reconcile this extremely restricted ontology with rational moral philosophy tend to be self contradictory. Thus he extols Daniel Wegner's book *The Illusion of Conscious Will* by saying (p. 224): "I think Wegner's account of conscious will is the

best I have ever seen", but then rejects Wegner's basic claim, and asserts the exact opposite, in a move he downplays as a mere *tactical* difference. He says that "Wegner thinks it is less misleading, more effective, to say that conscious will is an illusion" but that he, Dennett, thinks the better 'tactic' is "to make *the same points* by saying that no, free will is not an illusion."

In order to construct a rational moral philosophy concordant with the precepts of classical physics Dennett wants to say that we are mechanistically deterministic beings that enjoy free will. This position requires a twisting of the usual meanings of words that does not seems to fly: even those who generally agree with him seem unable to go along with these tamperings with normal meanings attached to the words he uses.

This commentary on Dennett's efforts to reconcile mechanistic determinism with rational moral philosophy is meant to emphasize that this is not a problem that can be justifiably claimed to have now been well resolved by philosophers. Books and books have been written, it is true, but great volumes of argumentation, followed by repeated re-argumentation, are insufficient, by themselves, and are, instead, a cause for serious skepticism and doubt.

Dennett asserts (p. 14) that his "fundamental perspective is naturalism, the idea that philosophical investigations are not superior to, or prior to, investigations of the natural sciences, but in partnership with those truth-seeking enterprises, and that the proper job for philosophers here is to clarify and unify the often warring perspectives into a single vision of the universe. That means welcoming the bounty of well-won scientific discoveries and theories…" Accordingly, he welcomes the offerings of (neo)Darwinism, but effectively rejects, or rather grossly misunderstands, the equally important offerings of quantum theory.

In his chapter 4 Dennett does consider the idea that the indeterminism of quantum theory might open the door to the entry of a free will that can aid in the construction of a rational moral philosophy. He poses the key question: "How can the indeterminism of quantum physics be harnessed to give us a clear, coherent picture of a human agent exercising this wonderful free will?" But he then

poses the question in an essentially different way: “How, exactly, could subatomic indeterminism yield free will?” Then, rather than considering quantum theory itself, he goes on at great length to discredit a model constructed by Robert Kane, who introduces a *random element* of indeterminism to break the absolute determinism of classical physics in the hope of thereby exploiting quantum theory to open the way to a satisfactory concept of free will. However, the replacement of conscious choices by random or whimsical choices certainly cannot provide a rational basis for morality. What is needed is not the injection into human behavior of pure whimsy. It is rather what quantum theory so beautifully and adequately supplies, namely the *suppression of chance and randomness* by choices stemming from a process arising, not from either mindless motions atoms or meaningless bits of noise, but rather from the wholeness, the meanings, and the physical efficacy of our conscious thoughts.

Later on (p. 223) Dennett asserts that he believes there is a morally important non-supernatural free will, but that it is “just not what you probably thought it was.” But what does the scientifically knowledgeable reader think a non-supernatural free will is?

Nature certainly contains non-supernatural thoughts and feelings: they are non-illusory *real* parts of the natural world. Both Dennett and quantum physics agree that they are causally efficacious, and are in some sense “free”. But it is Dennett’s own *pre-judgment*, not contemporary science, that insists that each of these components of a stream of consciousness is *identical* to some mechanistic material processes. It is his rigid commitment to the ideology of classical physics that is the basic source of the difficulties Dennett encounters: it is that choice that foists upon him the impossible task of showing that mechanically deterministic automata possess conscious free wills that can underlie rational moral philosophy.

The bounty offered by quantum theory is not the introduction of meaningless whimsy. It is the introduction of immaterial causes. The indeterminism introduce by quantum theory comes in two forms. One consists of the random “choices on the part of nature”. These conform to certain statistical laws linked to the quantum theoretical structure that *replaces* the material structure postulated by classical physics. The other form of indeterminism stems from a physically

efficacious conscious will that is “free” in the sense that it is not fixed by any *yet-known, or mechanical,* laws These consciousness-based choices are perhaps the most *real* elements in quantum theory: the whole theory is built around these conscious “free” choices on the part of human agents. They are the *least supernatural* element in science. Mathematical representations, formed in our endlessly creative imaginations, of some assumed-to-exist objective world ‘out there’ can come and go, and can morph in strange ways. But our choices about how we act will probably remain a stable element of science, because we can, by acting in diverse ways and observing the feedbacks, discover many more aspects of nature than we can by mere passive witnessings. Our consciously activated probing actions are fundaments of human life, and of both science in general and quantum theory in particular.

Dennett correctly poses the key question: “How can the indeterminism of quantum physics be harnessed to give us a clear, coherent picture of a human agent exercising this wonderful free will?” The basic thrust of the present book has been to give a specific answer to this question, together with the background needed to understand that answer. The answer in brief is that the quantum Zeno effect provides a nonlocal consciousness-driven process that permits a person’s conscious volitional choices to exercise significant control over the activities of his or her own brain.

Left unaddressed, however, was the task of specifying the causal origin of these conscious choices.

Pursuit of this important question takes us beyond the realm of contemporary physical theory. I have taken great pains to keep the content of this book securely in line with orthodox contemporary physics, by which I mean the main-line Copenhagen interpretation, extend by the work of John von Neumann to cover the activities of human brains

But now, in this final appendix, with a strong advisory that I am venturing into a realm of speculative philosophy, where contemporary physics provides no orthodoxy, I shall tackle this remaining issue in a way that appeals to my own personal intuitions. No claim is made that these speculations represent the best thinking of physicists or of

philosophers. They are simply my own best guesses, for what they are worth, unsupported by any new confirming data. Still, educated guesses, *compatible* with what is already known, are an important element of science.

The issue here is the causes of the conscious choices made by the human agents about how they will act. My own prejudice is that nothing happens without a sufficient cause: there is no pure whimsy and, moreover, the next step in basic science will embrace the idea of a process of a "coming into being" that converts "potentialities" into "actualities", as suggested by Heisenberg. But nothing definite can come into being simply "out of the blue". So these choices must have some reason to be what they turn out to be, instead of something else. Indeed, we seem to be at least dimly aware of some of our reasons for making the choices that we make, and psychiatrists---and more recently neuroscientists---have made claims about such cause that are not completely at odds with empirical evidence. How can such causes be understood in a way concordant with quantum theory?

I begin with the supposition that the "conscious choices on the part of the human agents" that quantum orthodoxy is built upon do indeed exist (or at least that a rationally coherent science-based understanding of the structure of human experience is possible in which these choices do exist) and that these volitional choices have necessary and sufficient causes, i.e., have reasons to be what they turn out to be, instead of something else. In other words, I am starting with orthodox contemporary quantum theory, and trying to complete it in a way that makes it part of an ontology that explains the nature of the factors that determine the "free choices" of orthodox contemporary physics.

I assume that our choices do not pop out of nothingness, and are in fact influenced by our prior experiences. This influence, or correlation, must be accounted for by a complete theory of nature. Quantum theory is not a complete theory *of nature*, because it does not explain the causes of our so-called "free choices".

The mathematical structure of quantum theory asserts that whereas the evolution of the quantum (i.e., physical) state of a system in

accordance with von Neumann's Process 2 is essentially mechanistic (i.e., locally deterministic), the interventions of Process 1 are essentially non-mechanistic (i.e., nonlocal): the effect of each such intervention acts instantaneously over an extended spatial region, and it can have, moreover, instantaneous (i.e., faster than light) subtle effects far away. Each Process 1 intervention represents an "observation" or "actualization of a particular potentiality from amongst the multitude specified by the prior quantum state".

The simplest way to impose a modicum of regularity on this structure is to suppose that there is a variety of elementary properties that can be combined in a variety of possible ways in a Process 1 event. Each such event "occurs" at some instant of time and covers some fixed region of three-dimensional space. Some of the elementary properties involved in such an event may be completely describable in terms of mathematical properties localized essentially at space-time points, whereas some elementary properties can combine to give, for example, the experiential quality of greenness, or the sorrow at the death of a beloved companion. These latter qualities are conceptualized in psychological or experiential terms. That is, certain elementary properties can combine to give experiences of the kind that populate our streams of conscious experiences.

Each Process 1 event, or happening, or actual occasion, is a gathering together, or prehension, of a generally large number of properties in some fixed region of space at some instant of time. Each such event fixes, or actualizes, in that region, a certain combination of properties that serve as potentialities for future events. Each future event will re-actualize, in a newly combined novel way, various properties---or parts of properties---previously actualized in certain prior events.

This general structure is very similar to the structure found in contemporary physics, where the dynamics can be understood as built out of a series of "scattering events", each of which gathers together and combines certain physical properties, drawn from a certain set of prior scattering events, and then serves as a source of properties that can be drawn upon by subsequent scattering events. One key difference between classical physics and quantum physics is that in the former the entire sequence of happenings is fixed by initial

conditions, whereas in the latter the *physical* constraints permit, *for any possible initial condition*, a huge blur of subsequent possibilities: the *physically described constraints* specify only a collection of subsequent future possibilities, each with a statistical weight. They do not fix what actually happens, or is experienced by observers. Another process, Process 1, is needed to specify what actually happens. It involves properties not entering into classical dynamics.

During the twentieth century theoretical physics has generated many important properties beyond the simple space-time and energy-momentum properties identified and used in classical physics. These new properties are “internal” properties of particles, which, however, can combine to form important macroscopic (i.e., large-scale) properties. Physical theories certainly need not be limited to the use of properties identified and applied in classical physics.

There are two big differences between the classical and quantum models. The first is nonlocality. The Process 1 reduction (or actualization) event is nonlocal, in the sense that it happens instantaneously over an extended region, and integrates the various properties brought into that region from certain prior events. Process 2, on the other hand, is a local process. However, it generates only “potentialities” for what might or can happen, not a definite fixing of what actually does happen. The *Potentialities*, like their classical counterparts, evolve according to local laws, but the *actualities* are determined by a process that is *not* locally deterministic: it is not “mechanical”, in my terminology.

The second big difference is this distinction between (1), the *potentialities*, which are governed---between the interventions of the Process 1 actions---by purely physical laws, and (2), the *actualities* which are created by the Process 1 interventions, which can involve experiential qualities. The physically described features governed by Process 2 specify only the *necessary* conditions for a possible event, in the sense of a statistical weighting of the various possibilities, whereas experiential or proto-experiential qualities enter into the *sufficient* conditions for the occurrence of some particular one of the many physically allowed events. In the classical theory the necessary and sufficient conditions for an actual event to occur are not separated in this way.

Each quantum Process 1 event draws upon past experiences as the grist from which it forms its own experiential nature. Our human experiences are *examples* of the experiential aspects of Process 1 events. But they are very complex examples, associated with very complex physical systems, which have evolved to take advantage of the existence and properties of this process. But Process 1 events can occur, and presumably do occur, also in conjunction with far simpler physical systems. In these physically simpler cases the "experiential" aspect can be something far simpler than anything we human beings would recognize and describe as an experience.

This brief sketch does not describe or constitute a physical theory: it is merely a general outline of how quantum theory might someday be extended from its present pragmatic form into a coherent idea of how Nature works. It is a rough blueprint for a future endeavor. I include it here only to dispel the idea that we must be dealing with some super-natural phenomena that are anthropocentrically tied to human beings, or to biological entities, or that "defies the laws of physics". We are dealing, rather, with a real aspect of nature that is neglected by, and excluded from, classical physics, but that constitutes an essential constituent of the actual world in which live and consciously act.

This loose sketch of a conceivable quantum ontology is in overall general agreement with the basic ideas of Alfred North Whitehead (1929/1978). Many physicists, including Abner Shimony (1993), Rudolf Haag (1996) and myself (Stapp, 1975, 1977, 1979), have noted the suitability of Whitehead's ideas as a foundation for an ontological completion of quantum theory. But the incorporation of such ontological speculations into empirically backed science awaits the arrival of secure pertinent data.

The morality that emerges from science will depend upon what science reveals about our connection to the universe that sustains us, and upon our judgments about how to orient ours lives in relation to it. But an appreciation of the established fact that the natural world enjoys a deep interconnectedness that goes far beyond what the precepts of classical physics can permit, and that the course of physical events can, according to contemporary physics, be influenced by the intervention of physically effective personal actions

that stem from our conscious reflections on the meanings of our actions, arms us with an image of ourselves that allows us to be far more significant parts of the scheme of things than the classical-physics-based image of ourselves as lonely bits of protoplasm, spawned by some freakish accidents of nature, and existing only because we serve as a vehicles for the survival of our genes. That impoverished self-image leads rationally to the meaningless of it all, and, inevitably, to the downward spiral of humanity into a society of machines that is even now being erected by zealous seekers of a simplified and more controllable social order that will bring our lives into alignment with the “established” classical-physics-based conception of what we human beings are.