

July 26, 2004

LBNL-55887

**Quantum Leaps in Philosophy of Mind:  
Reply to Bourget's Critique**

Henry P. Stapp  
Lawrence Berkeley Laboratory  
University of California  
Berkeley, California 94720

**Abstract.** *David Bourget has raised some conceptual and technical objections to my development of von Neumann's treatment of the Copenhagen idea that the purely physical process described by the Schrödinger equation must be supplemented by a psychophysical process called the choice of the experiment by Bohr and Process 1 by von Neumann. I answer here each of Bourget's objections.*

I answer Bourget's objections in the order he presents them.

Bourget's first objection involves juxtaposing a quote from Stapp (1997) and a quote from Stapp (1993). The first of these asserts that the uncertainties in the process of exocytosis arising from the quantum uncertainties in the motions of calcium ions inside nerve terminals entail that "the brain must evolve into an amorphous superposition of states corresponding to a continuum of different macroscopic behaviours." Then from my 1993 book he extracts the time  $200\mu\text{s}$  for the transit time of calcium ions inside nerve terminals from an ion channel to a trigger site for the release of the contents of a vesicle of neurotransmitter.

Bourget suggests that I equate this  $200\mu\text{s}$  transit time for a calcium ions inside a nerve terminal with the time for the macroscopic brain state to evolve. I make no such identification. The time required for the uncertainties at the synaptic level to develop into macroscopic uncertainties involving large numbers of neurons is surely very much larger than the transit times for individual calcium ions inside tiny nerve terminals.

Bourget next suggests some wavering of my understanding of the role of "will". He notes that I say (Stapp 1997, p.187) "...I stay strictly within the bounds of contemporary orthodox science in accepting the quantum statistical rules as primitive elements of our basic theory." Later (p.191) I say "In this paper I have adhered to the orthodox position that the quantum selection process is bound by the quantum statistical rules of contemporary quantum mechanics. Any suggestion that this law fails in certain cases should be supported by powerful data or reasoning, for a failure of this law would open a Pandora's box of theoretical difficulties and uncertainties." Also, in connection with, and in contrast

to, the approach of Eccles, I say (Stapp, 1993, p.79) ‘This proposed solution requires no ... distortion of the laws of physics.’ These statements make clear the fact that I am thoroughly committed to strict adherence to the orthodox statistical rules, in contrast to Eccles’s suggestion that these laws are biased by mental effort.

Bourget suggests there is difficulty reconciling this strict commitment to the quantum rules with my assertion (Stapp, 1993, p102)

The basic idea of the present psychophysical theory is to identify the *selection* of one of these mutually exclusive self-sustaining patterns of neural excitations as the image in the physical world, as represented by quantum theory, of a creative act from the realm of human consciousness.

The reconciliation between the *willful* and the *statistically lawful* aspects of this selection rests on the fact that two very different processes contribute to this selection. The first is von Neumann’s Process 1 (also called the “Heisenberg Choice”) and the second is what Dirac called “a choice on the part of nature”, and which I call the “Dirac Choice” or “Process 3.”

The pertinent consequences of this distinction were explained in detail in Stapp(1999), and they are a key part of my theory. I emphasized there that the rules that determine *when* Process 1 events occur, and *what* the associated projection operators P are, are not determined by the laws of quantum theory. This opens a “dynamical gap” that allows mental effort to influence brain activity. My general assumption, which is in general accord with Copenhagen and von Neumann quantum theory, is that the choices of *when* the Process 1 occurs, and *what* the associated projection operators P are, are determined *jointly* by mental and physical aspects of the mind-brain system, and that the quantum statistical laws, *which govern Process 3*, are strictly observed. Thus the potential conflict described by Bourget between “willful selection” and “statistical rules” is neatly evaded: freewill comes in through the *not-statistically-controlled* Process 1, not through the statistically controlled Process 3.

The detailed psychophysical rules that are presumed to fix or determine the projection operator P (which defines the empirical question) are not yet known, and I have made no basic commitment concerning the form of those rules. But I do postulate that mental effort can *increase the rapidity* of the Process 1 events. This postulate is the basis of my explanation of the effect of mind upon brain. Conjoined with the quantum Zeno effect, it can explain a very large number of empirical findings in psychology and neuroscience that pertain to effort and will (Schwartz, Stapp, & Beauregard, 2003, 2004.)(Stapp, 1999, 2004a,b,c)

Bourget asserts that “In his 1999 account he hints that the will can determine what measurement occurs, but he retracts this suggestion in his more recent

writings. Instead, he holds that the empirical question is always determined by the wave function of the brain at any time. “

Neither part of that claim is accurate. I assumed from the start, and still assume, that this choice is in principle determined jointly by mind and brain.

In Stapp(2001) and Stapp(2003) I gave, “A Simple Dynamical Model” in which the choice of the question (i.e., of the projection operator P) was determined solely by the brain (i.e., by the wave function of the brain), and that the mind determined only the rapidity of the (Process 1) events. But that was specified to be a simple model. I did not specify that *in general* “the empirical question is always determined by the wave function of the brain”, although that is an interesting possibility.

In that simple model a person will indeed, as Bourget notes, have to wait to raise his finger until the brain presents this possibility of raising the finger as the possibility he must either choose to consent to or not consent to. This is in accord with William James’s account of willful action, as I emphasized in Stapp 2003, 2004a,b,c. I believe these detailed accounts obviate the need for the “better picture” sought by Bourget. In any case, he admits that there is no *principled* difficulty at this point.

Bourget’s sections V and VI are not objections to my approach. they seem merely to explain some features of it.

In section VII Bourget asks whether in my theory the “collapses” are real. My answer is that when I pass from the purely pragmatic (e.g., Copenhagen) stance to an ontological stance I take the psychophysical Process 1 collapses to be real. The Process 3 events can also be considered real.

He then asks a key question: Can the conditions resulting in a ...collapse be described quantum mechanically? This question is linked to the question: What sort of physical systems support collapses?

My *partial* answer to the second question is that conscious human brains are physical systems that support collapses. This partial answer arises from the fact that quantum theory, in its original pragmatic Copenhagen form, was a set of rules that allowed human scientists to make computations connecting what they knew on the basis of past experiences to what they were likely to experience in the future under the various possible experimental conditions between which they were free to choose. The data supporting quantum theory and classical physics are all data of that kind. Venturing outside the domain based on human experience lands one in a realm of speculative philosophy that lies beyond narrowly construed science.

The data of psychology and (human-subject) neuroscience can be encompassed by staying in the more limited realm. I believe it prudent to focus first on this huge field before jumping into the more speculative arena. Still, I would want to grant to all biological systems some at least primitive form of mind, together with the capacity to effect collapses that goes with it. But I judge it premature to venture further into that speculative domain at this time

Bourget also asks the key question: Can the conditions resulting in a ...collapse be described quantum mechanically?

I guess he means the necessary and sufficient conditions that specify both when and where the collapse will occur and what that collapse will be---i.e., the value of the projection operator  $P$  specified by the Process 1---and which answer the Process 3 collapse specifies.

In regard to this question my theory, as presently defined, could go either way.

If the answer to Bourget's query is 'Yes', then the theory would be basically a deterministic physical theory, although not a local one. (The projection operators  $P$  must in any case be nonlocal, because an operator  $P$  acting at a point would dump an infinite amount of energy into the brain, causing it to explode.) This form of the theory, although deterministic within the physical realm alone, would not be equivalent to (basically local) classical physics. The evolution of the brain would be strongly influenced by the nonlocal process associated with conscious experiences via the collapses associated with these experiences.

In this model it would not be clear whether, for example, the increased rapidity of Process 1 events asserted to be associated with conscious effort should be considered to be *caused* by this (very real) conscious effort, or by some presumed-to-exist complex property of the physical state that could be supposed to invariably accompany such an effort. But, certainly, for practical considerations the conscious effort is the more useful candidate, because we know it exists, and persons can issue reports about it that can be used in our scientific studies, whereas its presumed physically described counterpart is both unknown to us at present, and it may remain forever inaccessible to us both empirically and also theoretically, even if it does exist. The physical determinist's contention that "someday" we shall be able to understand the presumed-to-exist purely physical determinants for the collapses is questionable not merely because an intensive seventy-five year effort has failed to find it, but, much more incisively, because the mathematical aspects of quantum theory are built upon the (Fourier-transform-based) uncertainty principle, which ties the theory naturally to probabilities, not to definite actualities. Something besides that quantum mathematical structure seems needed to fix actualities. This consideration tends to support to the opposing option 'No'.

If the answer to Bourget's query is 'No', then which individual Process 1 event occurs would *not* be fixed by the immediately prior instantaneously existing physical state *alone*. According to standard quantum ideas, the evolving quantum state does not contain enough information to fix the actually experienced flow of events themselves: something else is needed to fix these streams of discrete psychophysical realities.

But the only other realities that quantum theory speaks of, or that we know of, are the streams of conscious events that constitute our mental lives. So the simplest proposal concordant with quantum ideas *is to take those realities, and the analogous ones associated with other "agents", as the "other contributors" to the determination of the flow of actual Process 1 events*. These realities, once created, as parts of streams of consciousness, are, at least in principle, available to help do the task that the evolving physical state seems logically ill-equipped to do by itself, namely to determine what actually happens.

There has always been a basic question lurking within the deterministic idea that what is going to happen next is fixed by what has already happened. This question is: How much of the past is needed?

All that is needed in classical mechanics is a thin temporal slice extending infinitesimally into the past. The classical physical description is in terms of variables that are *continuous in time*. The quantum description has an analogous continuously varying part, namely the evolving physical state, and it might be reasonable to assume that, as in the classical case, only an infinitesimally thin temporal slice of this *physical* structure contributes to the causal structure. However, the *psychological* aspects are discrete: in the von Neumann formulation they come into existence in conjunction with an instantaneously existing physical part, which then immediately moves continuously onward. But the psychological aspects, being discrete: cannot evolve continuously.

This raises a problem: If discrete psychological realities from the past collaborate with the continuously evolving instantaneous physical state 'Now' to produce the current psycho-physical event, then in what part of the past are the space-time locations of the creations of causally contributing psychological factors confined? An infinitesimal slice would contain no such events at all!

Since each new experience *at least seems to be* in part a re-living of re-assembled fragments of old experience, some dating back to childhood, it would seem arbitrary to restrict the set of psychological elements that can causally contribute to the next actual event to psychological elements created in some thin but finite immediately preceding temporal slice. It may be unnecessary to impose such a restriction, because the strict implementation of all the quantum law would provide protection against wildly acausal behaviour. So, at least initially, I impose no condition on the locations where the causally relevant psychological elements were created.

These suggestions provide an ontological scaffolding in which are situated the psychologically described realities that are supposed to cooperate with the evolving quantum state to produce the actual sequence of psychophysical Process 1 events that constitute the cornerstone of von Neumann's formulation of quantum theory. This ontological proposal goes well beyond what the pragmatic Copenhagen philosophy encompasses, but is in general alignment with the Copenhagen idea that the evolving wave function does not suffice to specify reality, and that human choices do enter into the flow of human events.

Bourget claims to show in subsection VII.1 that I must accept both of the two incompatible options, and that the second possibility is not compatible with the 'freewill' it is supposed to accommodate. However, due to ambiguities in the meaning of "correlations" and of "freewill", his arguments do not logically prevent me from rejecting the first (physically deterministic) option described above in favor of the second (psychophysically deterministic) option, which I have spelled out in order to expose the non-generality of his arguments. The compatibility of this second option with "freewill" is permitted because (for me) freewill does not mean choices that pop out of the blue, determined by nothing at all. It means choices determined not solely by mindless mechanical processes, but rather, in part, by an evaluative process that is described essentially correctly by the psychological descriptions of what we feel or experience is happening.

Actually, even the first option, 'Yes', although physically deterministic, might better be described as psychophysically deterministic, and compatible with freewill. The point is that even if there is a strict isomorphism between each psychological reality and its physical brain counterpart, so that "in principle" the former can be eliminated in favor of the latter, the laws that determine what actually occurs might be directly expressed in terms of psychological variables, instead of their complex physical counterparts. That is, the evaluative process that selects one particular Process 1 event from the continuum of possibilities may deal directly with the psychological level of description. The process that determines which Process 1 event actually occurs would then quite properly be called "psychophysical", and the rooting of the evaluative process of selection in psychologically described qualities would qualify the process to be characterized as "freewill". The process would be non-mechanical, in the sense of not being controlled purely by local atomic-level processes.

The foregoing discussions reveal significant departures of my position from ones ascribed to me, and criticized, by Bourget in section VII.2. For example, I do not claim that Process 1 *cannot* be given a quantum physical explanation, Rather, I argue that once classical physics is replaced by quantum physics there is natural room for non localized psychological realities *per se* to enter significantly into an evaluative process that determines what an agent chooses to do.

Bourget raises the issue of solipsism. Quantum theory is basically a system that connects various psychologically described realities to each other by means of mathematical concepts. Take away all such realities and nothing remains. Thus the fundamentally psychophysical nature of pragmatic quantum theory should be carried over to its ontologicalization. A mindless quantum universe, unlike a mindless classical universe, is a non sequitur.

Bourget objects to my formulation of quantum theory on the basis of the fact that it is not equivalent to certain other collapse interpretations. That is not a valid objection. My formulation is based directly upon, and rationally extends, the formulation of John von Neumann.

The objection Bourget raises in Section VIII is technical, and a nonstarter. Its basis is the empirical fact that “two events that occasion simple conscious experiences must be separated by approximately 20 to 50 ms to be perceived as separate in time” However, the events in a rapid sequence generated by a willful effort are not perceived as separate in time. They all run together, and their rapidity is associated with the feeling of effort. I have never found any reason why the temporal separation between the events experienced in this run-on way could not be a fraction of 1ms, if their rapidity is controlled by conscious effort.

Bourget suggests that it is doubtful that there is a template for action “for each of the huge (if not infinite) number of actions that we can elect to perform”. He uses the phrase “quantum switch”, but in the context of my theory it should be “template for action”. I see no problem with the idea that each action that a person might choose to perform has a corresponding template for action, which is an executive pattern of brain activity that guides the action.

In section IX Bourget says “that superpositions of brain macrostates can disappear only through actual events that encompass both the brain and its environment”. In the von Neumann formalism that I use the “states” are represented by ‘density operators’ rather than by state *vectors*, and the state of a person’s brain is represented by partially tracing the state of the universe over all variables other than those of his brain. The operators  $P$  act as identity operators on all the non-brain variables. The Process 1 “collapse” is represented by  $S \rightarrow PSP + (1-P)S(1-P)$ , where the  $S$  can be taken to be either the state of the universe, or the state of the brain. The key point is that  $P$  acts nontrivially only in the subspace associated with the brain.

Bourget asks: “Why are only certain subsystems within the scope of an actual event relevant to conscious experience?” In my theory what is actualized in a Process 1 event is precisely a template for the intended action. The physical pattern of neurological activity that constitutes the template for action is certainly relevant to the conscious intention that it physically implements.

In my terminology thoughts are conscious thoughts. The unconscious Process 2 controlled by the Schrödinger equation does most of the analytical and constructive work, with conscious Process 1 selections providing only top-level direction.

I have answered cleanly all of Bourget's objections.

**Acknowledgement.** This work was supported by the Director, Office of Science, Office of High Energy and Nuclear Physics, Division of High Energy Physics, of the U.S. Department of Energy under Contract DE-AC03-76SF00098.

### **References.**

Bourget, D. (2004), Quantum leaps in philosophy of mind: A critique of Stapp's theory. *J. Consciousness Studies*, December 2004 (?).

Schwartz, J., Stapp, H. & Beauregard, M (2003). The volitional influence of the mind on the brain, with special reference to emotional self regulation. In M. Beauregard (Ed.), *Consciousness, Emotional Self-Regulation and the Brain*. [Advances in Consciousness Research Series]. Amsterdam & New York: John Benjamins.

Schwartz, J., Stapp, H. & Beauregard, M. (2004), Quantum Physics in Neuroscience and Psychology: A Neurophysical Model of Mind/Brain Interaction. [Submitted for Publication]  
[<http://www-physics.lbl.gov/~stapp/stappfiles.html>]

Stapp, H.P. (1993), *Mind, Matter, and Quantum Mechanics*, First Edition (Spinger-Verlag, Heidelberg, Berlin).

Stapp, H.P. (1997), 'Science of Consciousness and the Hard Problem'. *The Journal of Mind and Behavior*. 18 (2-3), 171-194.

Stapp, H.P. (1999), 'Attention, Intention, and Will in Quantum Physics, *Journal of Consciousness Studies*, 6 (8-9), 143-164.

Stapp, H.P. (2001), Quantum Theory and the Role of Mind in Nature. *Foundations of Physics*, 31, 1465-1499.

Stapp, H.P. (2004a), *Mind, Matter, and Quantum Mechanics*, Ch 12. Second Edition (Spinger-Verlag, Heidelberg, Berlin).



Stapp, H. P. (2004b), Quantum Approaches to Consciousness, *in Cambridge Handbook of Consciousness*, eds. M. Moskovitch & P. Zelazo. (To appear in 2005, Cambridge Univ. Press, Cambridge U. K.)  
[<http://www-physics.lbl.gov/~stapp/stappfiles.html>]

Stapp, H.P. (2004c), *The Mindful Universe*, [Book in Prep.]  
[<http://www-physics.lbl.gov/~stapp/stappfiles.html>]