

Comments on “The Illusion of Conscious Will” by Daniel. M. Wegner.
[M.I.T. Press, 2002] [Wegner is Professor of Psychology at Harvard]

This book aims to refute the notion that Conscious Will causes human actions. The question is: Do Wegner’s arguments, or the supporting data, overturn the von Neumann/Stapp theory of consciousness, which claims to show, within contemporary physical theory, how Conscious Will *can* cause human actions?

Wegner’s Chapter 1 begins with the statement:

“It usually seems that we consciously will our voluntary actions, but this is an illusion.”

That assertion appears to be a clear statement of what Wegner intends to demonstrate.

Earlier, in the preface, he states that “we need to understand how conscious will might be an illusion, a feeling that comes and goes independent of any causal relationship between our thoughts and our actions.” This statement clouds the issue, because “our thoughts” would seem to encompass the feeling of conscious will.

This lack of clarity is reinforced by saying [in the preface] that “the experience of conscious will is created by the mind and the brain just as human actions themselves are created by the mind and the brain.”

But if human actions are created *partly by the mind*, then the feeling that those actions are caused partly by experience of conscious willing could be illusory.

Later on in the preface we find:

“If psychological and neural mechanisms are responsible for human behavior then why does it feel that we are consciously causing the things we do?”

But admitting that “psychological mechanisms are [*partly*] responsible for human behavior” seems to be admitting that our consciousness is *part* of the cause of the things we do.

Wegner’s intent could be expressed by his statement [p.3]:

“One might assume that the *experience* of conscious willing an action and the *causation* of the action by the person’s conscious mind are the same thing, As it turns out, however, they are entirely distinct, and the tendency to confuse them is the source of the illusion of conscious will that this book is about.”

But how can anyone confuse a mere “experience of conscious willing an action” with the factual issue of whether or not that experience is actually a link in a causal chain that gives rise to that action?

I suspect/believe that Wegner’s intent is to argue that: (1), the stream of consciousness is a *fictional story* that the brain creates, *along side* its generation of the physical actions; and (2), the elements of this tale lie outside the causal chain that gives rise to the actions; and (3), this invention by the brain is useful because it creates a simple understanding that helps us with life, in spite of its fallaciousness.

In any case, the questions in need of addressing are: Does the empirical evidence cited by Wegner entail that a person’s stream of consciousness cannot causally influence that person’s physical actions? To what extent are Wegner’s arguments predicated implicitly on the deterministic concepts of classical physics, which are known to be basically false, and, in particular, inadequate for the treatment of brain processes that depends of the dynamics of the motions of ions. Do Wegner’s arguments, or the empirical evidence upon which they are based, create any difficulties for the von Neumann’s formulation of quantum theory, and the way that it explains the mechanism that allows a person’s “experiences of conscious willing” to influence that person’s physical actions?

The answers to these three questions are ‘No’, ‘Yes,’ ‘No.’

The key items of evidence that Wegner cites, were already carefully examined before the vN/S quantum account was proposed. Wegner's arguments and data, are not essentially new. His approach is basically that of epiphenomenalism, which suffers from oft-discussed problems.**

**

William James (1890: 138) argued against the possibility of epiphenomenal consciousness, by observing that "*The particulars of the distribution of consciousness, so far as we know them, points to its being efficacious.*" He noted that consciousness seems to be "an organ, superadded to the other organs which maintain the animal in its struggle for existence; and the presumption of course is that it helps him in some way in this struggle, just as they do. But it cannot help him without being in some way efficacious and influencing the course of his bodily history." James said that the study described in his book "will show us that consciousness is at all times primarily a *selecting agency.*" It is present when choices must be made between different possible courses of action. He further mentioned that "It is to my mind quite inconceivable that consciousness should have *nothing to do* with a business to which it so faithfully attends."(1890: 136)

If consciousness has no effect upon the physical world then what keeps a person's mental world aligned with his physical situation: what keeps his pleasures in general alignment with actions that benefit him, and pains in general correspondence with things that damage him, if pleasure and pain have no effect at all upon his actions?

Searle says [in the J. of Conscious Studies] that psychological indeterminism with neurobiological determinism---

"is intellectually unsatisfying because it is a modified form of epiphenomenalism. It says that the psychological processes of decision making really do not matter. The entire process is deterministic at the bottom level, and the idea that the top level has an element of freedom is simply a systematic illusion. ... The bodily

movements would be exactly the same regardless of how these processes occurred.

“Maybe that is how it will turn out, but if so the hypothesis seems to me to run against everything we know about evolution. It would have the consequence that the incredibly elaborate, complex, sensitive and ---above all---biologically expensive system of human and animal conscious rational decision-making would actually make no difference whatever to the life and survival of the organism. Epiphenomenalism is a possible thesis, but it is absolutely incredible, and if we seriously accept it, it would make a change in our world view, that is, in our conception of our relations to the world, more radical than any previous change, including the Copernican Revolution, Einsteinian relativity theory and quantum theory.”

The chief “difficulties” that Wegner finds with the idea that “we conscious will our voluntary actions” are, first, the Libet data, which are beautifully accounted for by quantum theory— indeed, that wonderful success was a key element of support for my quantum approach to consciousness – and, second, the fact that most actions occur automatically, without the causal intervention of conscious thought. That feature is also well accommodated. Hence the arguments of Wegner, and the data he cites, do not undermine the quantum theory of consciousness: in fact the crucial Libet data that he cites as a problem for the idea of efficacious consciousness actually provide substantial support for the quantum approach, as will be discussed below.

Wegner stresses “scientific psychology” and “empirical will,” His emphasis is on “objective data.” But positivistic philosophy is too restrictive for science: one must allow theoretical concepts that are not objectively defined, so long as useful objective conclusions eventually come out. Thus psychological variables, such as beliefs, desires, efforts can be introduced, and so can the concept that a person’s feeling of conscious willing can affect that person’s actions, provided the rules that govern this connection are well defined in a

way that leads to important contributions to scientific practice and prediction.

The question thus arises: How can a rational relationship be defined that connects the “world of conscious experiences,” which contains the empirical data, and the “world of microscopic physical variables,” which houses the theory.

Part of the answer is that some of the critical experiential variables are experiences described in terms of macroscopic geometric variables: experiences of *macroscopic* locations, motions, and shapes. The *physical* variables are *microscopic idealizations* of these macroscopic geometric concepts. This rational connection via microscopic idealization provides a logical connection between the two kinds of description upon which scientific practice is based. It allows the experiential aspect to be rationally connected to the quantum mathematics that describes the “physical” (i.e., microscopic space-time) aspects.

Von Neumann/Stapp quantum theory explains also, as we shall see below, how the psychologically/experientially described feeling of “effort,” or “will” is tied to the physical dynamics.

One of Wegner’s main arguments is based on his Theory of Apparent Mental Causation, which says:

“People experience conscious will when they interpret their own thought as the cause of their action.”

Granting that this is true we may ask: Does this effect support the idea that conscious will is non-efficacious?

The answer is ‘No.’

Only if experiences ARE efficacious is there a good reason for the Theory of Apparent Mental Causation to hold!

If our experiences do causally influence our actions in the way that they seem to do then these experiences are surely needed for

successful living. But how do we learn how to USE this power of our minds to affect our actions?

Look at any health infant and you see the answer: incessant struggle to find out 'which effort does what.' Success in this venture requires the capacity to recognize those relationships between efforts and feedbacks that *could be consequences of those efforts*. The validity of the Theory of Apparent Mental Causation has, therefore, positive survival value IF the experience of conscious will actually has causal efficacy, but it has no survival value if the experience of conscious will has no causal efficacy. So I would argue that the evidence accumulated by Wegner (and Wheatly) for the validity of the Theory of Apparent Mental Causation is evidence IN FAVOR OF True Mental Causation, contrary to the converse interpretation given by Wegner. A process that makes 'efforts that COULD have experiential causal feedbacks' *seem* effective can help us to discover, by trial and error, those particular efforts that ACTUALLY DO have experiential causal feedbacks.

The Libet Data.

Because Wegner and I both stress the importance of the Libet data, it may be helpful to review here the quantum treatment of this data, in order to show how and why, in quantum theory, the **experiential cause** of "the action of raising the finger" comes AFTER the onset of the readiness potential, in direct violation of the *classical* formulation of the idea that cause always precedes its effect.

There is also in quantum theory a strict prohibition of effects preceding cause. But in the quantum case the effects so limited are only USABLE effects: effects that can be CONTROLLED by the human choice maker. The readiness potential is not controllable in this required way. This is very important, so let me explain it in more detail.

Quantum theory is based on Heisenberg's seminal discovery that the empirical facts (many of which are logically incompatible with the general precepts of classical physics) can be described by a new theory, quantum theory, which can be constructed by replacing the

“numbers” in classical physics by “actions” (operators). [The ordering of the *numbers* in a product does not matter, but the order in which *actions/operations* are performed does matter.]

Quantum theory is built upon a postulated correspondence between certain actions/operators in the mathematical theory and associated human experiences. Each such action is represented by a “projection operator” P , which satisfies $PP=P$. [The double action PP of a projection operator P has the same effect as a single action P .] If the experience is labeled by ‘e’ then the associated projection operator is represented by $P(e)$. In von Neumann’s formulation of quantum theory this operator $P(e)$ acts upon the state of the brain of the observer/participant/agent and specifies the neural correlate of the experience ‘e.’ The operators $P(e)$ play an absolutely key and essential role in the dynamics of the brain of a conscious human being.

The (quantum) state S of a system is an action/operator, called “the statistical operator” [or “the density matrix.”] It specifies the statistical weight [probability] of every projection operator P associated with that system. The formula for the statistical weight of P in the state S is:

$$\langle P \rangle = \text{Trace } PSP / \text{Trace } S.$$

If A is an action/operator then, by definition, $\text{Trace } A$ is the number generated when A acts back around on itself, like a snake biting its own tail. The connection of the mathematical formulas to measurable numbers is always given by this Trace operation.

Quantum dynamics is built upon these operators $P(e)$ and S , and on *two kinds of choices*. The first kind of choice is made by the experimenter/observer/participant. It is called “Process 1” by von Neumann. I have also called it the “Heisenberg choice,” because its crucial role in quantum theory was strongly stressed by Heisenberg.

The other kind of choice was called by Dirac “a choice on the part of nature.” I have called it “The Dirac Choice.” I also call it “Process 3” to distinguish it from von Neumann’s Process 1 and Process 2. [Process 2 is the quantum analog of Newton’s classical equation’s of motion,

and is obtained by replacing classical numbers by corresponding quantum operators.]

In classical physics there is just one dynamical process, namely the classical approximation to the quantum Process 2. But quantum theory has two additional processes, one involving a choice made by a conscious participant/agent/observer about how he will act, and one made by nature about how she will respond to the agent's choice.

The general theory does not specify what the agents's choice will be. In atomic physics the agent's choice is treated as a free variable that is fixed by the aims of the experimenter/participant. These aims are considered to lie outside the realm of atomic physics. They are to be covered by neuro-psychology, and are presumably determined by some combination of neurological-physical and psychological-experiential processes that enter into quantum neurodynamics. Figuring out exactly what this combination is, from a detailed analysis of the psycho-neurological data, is the task of neuroscience, psychology, and physics, working together. Only the general overall quantum framework was provided by von Neumann.

To provide to a preliminary general orientation, I have proposed a simple model for the participant's choice.

The state $S(t)$ of the participant's body-brain is defined by taking the "partial trace (over all *other* degrees of freedom in the universe)" of the state of the universe at time t (in the rest frame of the cosmic background radiation.)

Then the projection operator $P(t)$ is defined to be that operator in the set $\{P(e)\}$ that maximizes

$\text{Trace } P(e)S(t)P(e)/\text{Trace } S(t)$.

This special $P(t)$ is the $P(e)$ that has at time t the greatest statistical weight.

As a first guess, I propose that a Process 1 event associated with $P(t)$ occurs whenever Trace $P(t)S(t)P(t)/\text{Trace } S(t)$ reaches a local (in time) maximum. This Process 1 puts to Nature the question: Does the quantum jump to the state $P(t)S(t)P(t)$ occur?

Notice that the timing and form of this event is then determined jointly by the physical side, from $S(t)$, and by the psychological side from $P(e)$. But it is determined, nevertheless, by a mathematical law: it is not coming from “out of the blue.” The Process 1 event changes $S(t)$ to $S'(t) = P(e)S(t)P(e) + P'(e)S(t)P'(e)$, where $P' = (1-P)$.

Then Nature’s choice, Process 3, occurs. It is a “quantum jump.” The State $S'(t)$ is reduced to $P(e)S(t)P(e)$ with probability $\text{Trace } P(e)S(t)P(e)/\text{Trace } S(t)$ or to $P'(e)S(t)P'(e)$ with probability $\text{Trace } P'(e)S(t)P'(e)/\text{Trace } S(t)$. If the chosen state is $P(e)S(t)P(e)$, then experience ‘e’ occurs: otherwise no experience occurs in conjunction with this event.

The “experience” could be low grade, with no significant involvement of any “consent to apply effort”. In this case the automatic aspect of the quantum process is a consequence of these partly deterministic and partly statistically deterministic equations.

The causal efficacy of *mind* is introduced by postulating that the experience ‘e’ associated with a ‘Yes’ choice on the part of Nature can *allow* a mind-based (i.e., experience-based) “consent to apply effort” to occur. This “mental consent” will trigger a sustained rapid sequence of Process 1 events with nearly identical projection operators P , even in the absence of the physically specified maximal condition. If the rapidity of these Process 1 events is sufficient then this sequence will activate the Quantum Zeno Effect, which will tend to hold the state $S(t)$ in the subspace defined by the almost-constant $P(t)$. That near-constancy can produce a large behavioral effect, as I have explained in other places: the finger will be raised by the dynamical effects initiated by the conscious consent.

That was just a quick review of things I have often said elsewhere. It allows us to get to the point at issue here, which is the causal role of mind in the structure of the Libet experiments.

The point, now, is that the original commitment by the subject to, say, “raise my finger within the next two minutes” will bring on a sequence of Process 1 events associated with “virtual finger raisings,” distributed over the next two minutes. Each of these “virtual finger raisings” will be preceded by its required readiness potential. But the probability that any one will be selected by Nature’s choice, process 3, is very small. So most of the virtual finger raisings will not be actualized: they will be obliterated by the “No” answer on the part of Nature, and never be actualized. But for some one of these Process 1 choices on the part of the participant to raise the finger “now,” nature will say “Yes.” If the mind-based “consent” is then given, then a rapid sequence of Process 1 events will be actualized, and this will cause the finger to rise.

But the mind-based consent that causes the finger to rise then occurs *after* the beginning of the build-up of the associated readiness potential. This readiness potential is actualized by Nature’s first ‘Yes’ answer. None of the “potential readiness potentials” associated with the ‘No’ answers to the earlier Process 1 will have been actualized. So the physical situation actualized by the ‘Yes’ answer at some time T to one of the Process 1 events will actualize a readiness potential that begins its build up before time T. The mental decision to consent (not veto) comes after T, and only if this consent is given will the Quantum Zeno Effect kick in and hold persistently in place the brain state needed to consciously raise the finger.

It might seem that this occurrence of the build up of the readiness potential *before* the mind-based consent that triggers the raising of the finger would violate causality requirements. But the computations of orthodox quantum theory show that this kind of precursor activity *cannot be controlled* in such a way as to, say, send a specified message backward in time. It is controlled in this case by Nature’s choice to say ‘Yes’ at time T, not before. *Given* this ‘Yes’ choice on the part of Nature the (human) agent is given the choice to consent or veto the rapid sequence that will cause the finger actually rise. This human choice to consent or veto, on the basis of his feelings, can be treated as a free variable. But one must take into account the fact that if the consent is given then Nature must choose, with specified

statistical weights, between the 'Yes' and 'No' answers to each of Process 1 questions in the ensuing rapid sequence. The somewhat fantastic, but absolutely secure, result is that the granting of the consent can directly and strongly influence whether or not the finger will rise, but will have no effect, on the average, on whether or not the precursor readiness potential appears: the fact that that RP appears was fixed already by the 'Yes' answer given at time T. Consequently, the occurrence of the RP is not controlled by the subsequent "free choice of whether or not to consent," and there is no conflict with the stringent causal requirements of the theory of relativity, which forbids sending controlled messages except via physical transfers of momentum energy. There can no such transfer backward in time (or outside the forward light cone) and hence no violation of the requirements of the theory relativity, even though the readiness potential appears before the mind-based consent that actually *causes* the finger to rise.

The point here is that within quantum theory the Libet data are perfectly compatible with the idea that our conscious will can strongly influence our physical actions. Hence those data provide no support for the idea that the causal efficacy of conscious will is an illusion. Rather, quantum theory provides a plausible scenario for how the Libet data meshes with the causal efficacy of conscious will.

The projection operators $P(e)$ are necessarily nonlocal operators: they grasp in a unified way an informational structure that can extend over a large part of the body-brain of the participant. This brings into the dynamics holistic features that are in principle beyond the reach of systems that operate according to the mechanical principles of classical physics. But they are in line with our perception and conception of ourselves as creatures that can consciously grasp as wholes complex informational structures, and can act on the basis of those graspings.

Is it not more rational to abandon the unsatisfactory notion of epiphenomenalism, and the known-to-be-false classical physical theory upon which it is based, than to reject contemporary physical theory and the natural explanation it provides for the dynamical origin of the causal efficacy of our conscious will.