We Are Not Walking Wave Functions.

A Response* To

"Quantum Mind and Social Science" by Alexander Wendt

by Matthew J. Donald

The Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE, Great Britain.

e-mail: mjd1014@cam.ac.uk

web site: http://people.bss.phy.cam.ac.uk/~mjd1014

A. Wendt, *Quantum Mind and Social Science* (Cambridge, 2015), 354 pp, ISBN987-1-107-44292-4.

Alexander Wendt raises many interesting questions in this book, but to get to the answers he wants, he relies on a misunderstanding of the nature of the quantum states of macroscopic objects. I have been thinking and writing for many years about the relationship between the problem of "the observer" in the interpretation of quantum theory and the nature of information processing in the human brain (Donald 1990). By this time therefore only a fairly special book could both interest me and have as a central claim "that we are walking wave functions" (Wendt, page 3, see also pages 97, 137, 181, 201, 283, 289, and 290). I shall argue below that this claim is something like $10^{9\times10^{27}}$ steps wide of the mark, and that the conclusions Wendt draws from his claim are correspondingly unfounded. Nevertheless, leaving aside quantum theory, Wendt has carefully studied a broad literature and he has commented on it with perceptive clarity. I believe that he has been too uncritical of some of what he has read, but he has diligently noted many of the counter-arguments, even when he has not been swayed by them. Again and again, I found myself thinking of questions like "Surely this is just an analogy?", "What about decoherence in the brain?", "How could biological quantum computation evolve", "How are individuals constituted in panpsychism?", and again and again I read on for a few more pages and found that Wendt had come across the question and had a comment on it and a reference for it. I can therefore imagine recommending this book to a physics student thinking about consciousness and wanting to investigate some ideas from outside the conventional literature.

To justify his appeal to quantum theory, Wendt refers particularly to work on "quantum decision theory", for example the work discussed in the book "Quantum Models of Cognition and Decision" by Busemeyer and Bruza (2012). However, Busemeyer and Bruza make no explicit claim to be modelling brain functioning using quantum mechanics, and at best seem just to be introducing some formal tools for

* submitted March 2016, published Journal for the Theory of Social Behaviour **48**, 157–161, (2018).

the calculation of context- or order- dependent probabilities. Whether or not these tools are genuinely useful, the idea that they provide explanations, for example, of the way that people tend to tell stories about a socially-concerned female college student who might or might not become a feminist or a bank teller strikes me as simply absurd. Of course there are ways in which quantum theory can model some logical fallacies. Explanations in physics, however, require not merely the ability to replicate a handful of numbers, but details of physical structures and derivations from first principles of the laws governing their behaviour.

After explaining why I believe that Wendt's conclusions do not have the foundation he wants, I shall also argue against his premise that a foundation in physics is necessary for such conclusions. In the course of showing that we are not walking wavefunctions, implicitly at least, I shall also challenge Wendt's quantum coherence basis for panpyschism and for solving the "binding problem", and indicate why Wendt's suggestions about human mind-to-mind non-local dialogue come across as simply ridiculous to someone with my background. I hope it will become clear why neural quantum computation, if it exists at all, has to be in very small scale, highly isolated systems (this is a position Wendt explicitly rejects, on page 97, as the "weak thesis"). With that sort of computation, the questions I want to ask become "Where are those systems?", "How does input and output work?", "How did the systems evolve?", and "What biological function do they play?". For the small-scale quantum biological systems that I mention below, there are plausible answers to these questions. However the most important question is one to which I have never seen a satisfactory answer. That question is "Why would an alternative method of computation in the brain make a fundamental difference to the nature of human reality?"

Quantum theory underpins our understanding of all of modern physics, from subatomic particles to neutron stars, by way of chemical reactions and electronic devices. At the heart of quantum theory is the idea of a quantum system (the universe, a higgs boson, a large collection of neutrons, a pair of colliding molecules, a cloud of mobile electrons in a lattice of ions) and the idea of a quantum state of such a system. There are two types of quantum state: "wavefunctions" also called "pure states", and "mixed states". In a first course on quantum theory, mixed states may well not be mentioned at all as, in mathematical terms, they are just a certain sort of collection of pure states and almost all of the most interesting things which happen in quantum theory happen can happen to pure states. Indeed many of those interesting things are washed out when pure states are collected into a mixed state. This is where I believe that Wendt has gone wrong, because if whole human beings do have quantum states (and Wendt and I agree that they do) then rather than being wavefunction states, they are mixed state collections of at least $10^{9\times10^{27}}$ wavefunctions and so all of the quantum magic on which Wendt depends, like entanglement and non-locality, are supressed by a correspondingly large factor.

Let's consider a simple argument on this point. By "simple" here, I mean that I would expect anyone who regularly attends seminars in a subject which uses quantum theory, to be able to follow this argument easily. Although almost everything about the interpretation of quantum theory remains up for debate, without technical knowledge or at least collaboration with someone with technical knowledge, it is easy to go wrong and one is unlikely to make an original contribution to the debate.

According to quantum statistical mechanics, there are good reasons to identify the measurable physical entropy S of a quantum system in a state σ with the von Neumann entropy $-k_B \operatorname{tr}(\sigma \log \sigma)$. Here k_B is Boltzmann's constant (1.38×10^{-23} J K^{-1}). For example, 75kg of water at body temperature has $S = 3.0 \times 10^5$ J K^{-1} and so, for such a system, $-\operatorname{tr}(\sigma \log \sigma) = \log N$ where N is around $10^{9\times 10^{27}}$. N is much bigger than the number of molecules in the water, it's actually a measure, in classical terms, of the number of different ways that those molecules might arrange themselves at the given temperature. What is relevant for our present purposes is that N is the absolute minimum of the number of pure states into which σ splits, and so it is a direct measure of how unlike a wave function σ is.

Here are the details of that part of the argument:

Any quantum state can be written as a sum (the "collection") of orthogonal pure states : $\sigma = \sum_{i=1}^{M} p_i |\psi_i\rangle\langle\psi_i|$ where $0 \leq p_i \leq 1$ for $i = 1, \ldots, M$ and $\sum_{i=1}^{M} p_i = 1$. Then $-\operatorname{tr}(\sigma \log \sigma) = -\sum_{i=1}^{M} p_i \log p_i \geq 0$.

If σ is a pure state (a "wave function"), then $M = 1$. But, in that case, necessarily $p_1 = 1$, and $-\operatorname{tr}(\sigma \log \sigma) = 0$. More generally, subject to the given constraints, $-\sum_{i=1}^{M} p_i \log p_i \leq \log M$ with the maximum attained exactly when $p_i = \frac{1}{M}$ for $i = 1, \ldots, M$. This is "easily" proved using, for example, the method of Lagrange multipliers.

But that means that if $-\operatorname{tr}(\sigma \log \sigma) = \log N$, then we have to have $M \geq N$.

This argument doesn't leave much room for escape. Equating von Neumann entropy to physical entropy does require that the system is one to which it is appropriate to apply thermodynamics. But in thermodynamic terms, walking humans (or their brains or their neurons) really are pretty much warm water, in particular because the information ("negentropy") that we require for our daily lives is thermodynamically negligible. Given the magnitude of N, or just of the inverse of k_B in macroscopic units, the details of temperature and of particle nature are irrelevant. S and $\log N$ are proportional to the volume of the system. For a single water molecule $\log N$ is around 8.8. A cubic nanometer of water contains about 33 molecules, $\log N$ is around 290 and N something like 10^{126}. All that matters is that there is a number of particles each with an albeit-limited range of largely-independent random thermal motions. It is becoming increasingly clear, on the other hand, that there are nanometer-scale biological systems in photosynthesis, in olfaction, and in avian detection of the Earth's magnetic field in which specific degrees of freedom are sufficiently isolated from ordinary thermal processes that, on suitable time scales, thermodynamics is irrelevant and quantum coherence may play a significant role (Huelga and Plenio 2013, Al-Khalili and McFadden 2014). Avian magnetoreception in particular seems, with timescales possibly up to hundreds of microseconds, to be pushing at the boundaries of our understanding of the limits on quantum coherence (Gauger et al. 2011). However, whenever multiple-particle molecular or ionic diffusion processes are involved, for example in transmission at neural synapses (measured in tens of nanometers), thermodynamics will surely be relevant. The complete human systems which Wendt discusses are of scales sufficiently large that it would be straightforward to measure their temperatures (for example with a hidden infra-red camera) without the slightest effect on function. They are certainly in thermal contact with their surroundings; at least assuming that they are breathing.

The central reason why elementary courses in quantum theory can ignore mixed states is that, in an isolated system, pure states remain pure. This however is not true for systems which are not isolated. At every scale, we are made up of systems interacting within systems all the way up to the level of the entire universe. We can imagine that the state of the universe is pure, although we have no way of knowing, but, if we have a large system in a pure state, then most states of most subsystems of that system will be mixed. Indeed, some very interesting and fairly recent work has given precise senses to the words "large" and "most" in the previous sentence and confirmed the relationship to similar ideas of largeness and likelihood in statistical mechanics (Popescu, Short, and Winter 2006).

Wendt is correct to assume that quantum strangeness, and in particular usable quantum non-locality, requires that states be pure (or close to pure). He wants to argue that what Einstein called "spooky action at a distance" provides a necessary harmony between people in everyday conversations. I do not understand what this supposed harmony would actually do for us if it existed, but I am confident that, in fact, it does not exist. Quantum computers depend on the manipulation of pure states, and the difficulty we have in building useful quantum computers shows how difficult it is to control such states; even in a laboratory with access to very low temperatures and the ability to isolate individual atoms from their surroundings to a quite remarkable extent. At almost all relevant scales, biological systems can be thought of as warm and wet; the particles within them are constantly jiggled. Human conversations are very large scale process on time scales vastly longer than the associated thermal vibrations. From the synaptic level up, thermal vibrations give rise to an unpredictability in precise neural outputs (Ribrault, Sekimoto, and Triller 2011). There is no way in which the unpredictabilities in the synapses of one person can be finely tuned to those in the synapses of an interlocutor. Also relevant to Wendt's assumptions about non-local interactions is recent work on what has come to be called "quantum monogamy". This places limits on the quantum interactions simultaneously possible between more than two systems (Seevinck 2009).

I was drawn to the study of the interpretation of quantum theory by just the sort of philosophical dissatisfactions which motivate Alexander Wendt. I am afraid however that I do not believe that the kind of answers he seeks can be obtained. Wendt makes it clear that he rejects my current position, but I shall conclude by sketching it anyway, because I do not see any alternative.

As I understand it, Wendt's fundamental premise is that we will only ever be able to get away from treating each other as objects if physics can justify the reality of human consciousness. Unfortunately, although physics is good at modelling reality, there is, as discussed by Wendt, quite a wide range of models, in particular in the interpretation of quantum theory. Indeed, the more we explore, the more it seems that we are free within the framework of modern physics to paint a wide range of quite radically different pictures of how reality might be and that it is not likely that anything will ever let us definitively decide between those pictures. Moreover, while my own view is that there may indeed be clues about the reality of consciousness deep in the mysteries of the interpretation of quantum theory (Donald 2014), I really don't see why taking such clues seriously should alter our humanity. People aren't going to stop coveting their neighbours' houses, let alone their neighbours' oil, just because they think that they share a quantum spark. I believe that free will, for example, is a matter of how we think about our actions, rather than the expression of a genuine physical freedom. In my opinion, choice depends on thinking, and thinking depends on what our brains do, not, say, on the deeper thoughts of an immaterial homunculus. Nevertheless, I believe in the moral reality of free will. We can explain our choices and we can justify our choices, but ultimately, whatever physics may say, we find ourselves living our choices. We choose to punish criminals, and, however cruel or unfair we may be, in so doing we are at least offering them the same dignity, in accepting that they chose their actions, that we require for ourselves. It takes an effort to remember that the moral choices we see cartoon characters making are made up, but, this is precisely because it is the stories that matter not the causal roots of the stories. In much the same way, we see ourselves making moral choices without any need to trace those choices to their causal roots.

If you need to consult a biology textbook, let alone a physics textbook, to know if love is real, then surely you have never been in love. Polemics is not argument, but I do not think there is any stronger counter to Wendt's premise. Wendt's quantum theory may be unfounded but, even we were at some level non-local quantum computers, physics would still provide us with a range of external pictures according to which we just are what we are (Nagel 1986, Sartre 1943). What we need in order to see ourselves not as objects is an internal picture. I do not believe that there is a rational foundation for morality, but I also do not believe that it needs one. Wendt's argument seems to say that "if physics tells us we are just automata then we have no reason to care about each other", to which my response is that "if we can care about whether or not we are automata then we can care, and that is all the starting point we need". Getting from being able to care, to being able to accept the carings of others is either what caring is for, or it's an anti-solipsistic leap in the dark. This might lead us to attempt a "rational" division of other intelligences into "us" and "not us". The division between "non-automata" and "automata" might seem the least repugnant of such divisions, but, if it exists at all, it would be one of the most difficult to establish (unlike, say, the division between carbon-based and silicon-based intelligences). After we have decided who "we" are, we, no doubt with the help of our social scientists, are ready for the hard work towards mutual agreement about what constitutes a good society and how we can achieve it. Of course, in principle, goal-oriented automata could also do this sort of work. And then they could tell us that they were glad they had co-operated and, by doing so, had produced a harmonious, happy, caring society. I do not believe that we can do any better.

References

Al-Khalili, J. and McFadden, J. (2014) *Life on the Edge: The Coming of Age of Quantum Biology* (Bantam).

Busemeyer, J.R. and Bruza, P.D. (2012) *Quantum Models of Cognition and Decision* (Cambridge).

Donald, M.J. (1990) "Quantum theory and the brain." *Proc. R. Soc. Lond.* **A 427**, 43–93.

Donald, M.J. (2014) "From quantum theory to philosophical idealism." An after-dinner talk I gave for a general audience, the transcipt of which is available here: http://people.bss.phy.cam.ac.uk/~mjd1014/postp.html

Gauger, E.M., Rieper, E., Morton, J.J.L, Benjamin, S.C., and Vedral, V. (2011) "Sustained quantum coherence and entanglement in the avian compass." *Physical Review Letters* **106**, 040503.

Huelga, S.F., and Plenio, M.B. (2013) "Vibrations, quanta, and biology" *Contemporary Physics* **54**,181–207. *arXiv:1307.3530*

Nagel, T. (1986) *The View from Nowhere* (Oxford).

Popescu, S., Short, A.J., and Winter, A. (2006) "Entanglement and the foundations of statistical mechanics." *Nature Physics* **2**, 754–758. *quant-ph/0511225*

Ribrault, C., Sekimoto, K., and Triller, A. (2011) "From the stochasticity of molecular processes to the variability of synaptic transmission." *Nature Reviews Neuroscience* **12**, 375–387.

Sartre, J.-P. (1943) *L'être et le néant, (Being and Nothingness)* (Gallimard).

Seevinck, M.P. (2010) "Monogamy of correlations vs. monogamy of entanglement." *Quantum Information Processing* **9**, 273-294. *arXiv:0908.1867*