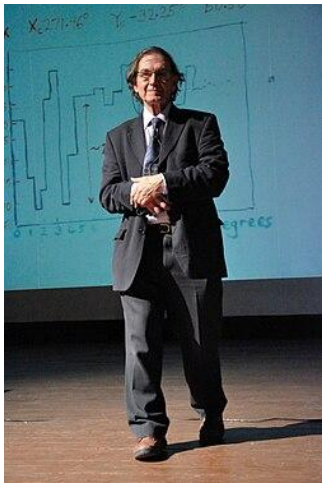
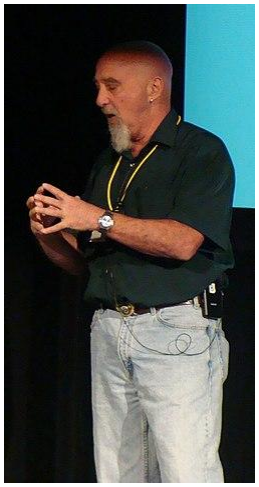
Orchestrated Objective Reduction





The founders of the theory: Roger Penrose and Stuart Hameroff, respectively

Orchestrated objective reduction (**Orch OR**) is a theory postulating that consciousness originates at the quantum level inside neurons (rather than being a product of neural connections). The mechanism is held to be a quantum process called objective reduction that is orchestrated by cellular structures called microtubules. It is proposed that the theory may answer the hard problem of consciousness and provide a mechanism for free will.[1] The hypothesis was first put forward in the early 1990s by Nobel laureate for physics Roger Penrose, and anaesthesiologist Stuart Hameroff. The hypothesis combines approaches from molecular biology, neuroscience, pharmacology, philosophy, quantum information theory, and quantum gravity.[2][3]

While some other theories assert that consciousness emerges as the complexity of the computations performed by cerebral neurons increases,[4][5] Orch OR posits that consciousness is based on non-computable quantum processing performed by qubits formed collectively on cellular microtubules, a process significantly amplified in the neurons. The qubits are based on oscillating dipoles forming superposed resonance rings in helical pathways throughout lattices of microtubules. The oscillations are either electric, due to charge separation from London forces, or magnetic, due to electron spin—and possibly also due to nuclear spins (that can remain isolated for longer periods) that occur in gigahertz, megahertz and kilohertz frequency ranges.[2][6] Orchestration refers to the hypothetical process by which connective proteins, such as microtubule-associated proteins (MAPs), influence or orchestrate qubit state reduction by modifying the spacetime-separation of their superimposed states.[7] The latter is based on Penrose's objective-collapse theory for interpreting quantum mechanics, which postulates the existence of an objective threshold governing the collapse of quantum states, related to the difference of the spacetime curvature of these states in the universe's fine-scale structure.[8]

Orchestrated objective reduction has been criticized from its inception by mathematicians, philosophers,[9][10][11][12][13] and scientists.[14][15][16] The criticism concentrated on three issues: Penrose's interpretation of Gödel's theorem; Penrose's abductive reasoning linking non-computability to quantum events; and the brain's unsuitability to host the quantum phenomena required by the theory, since it is considered too "warm, wet and noisy" to avoid decoherence.

Background

Further information: Penrose–Lucas argument

In 1931, mathematician and logician Kurt Gödel proved that any effectively generated theory capable of proving basic arithmetic cannot be both consistent and complete. In other words, a mathematically sound theory lacks the means to prove itself.[17] In his first book concerning consciousness, *The Emperor's New Mind* (1989), Roger Penrose argued that equivalent statements to "Gödel-type propositions" had recently been put forward.[18]

Partially in response to Gödel's argument, the Penrose–Lucas argument leaves the question of the physical basis of non-computable behaviour open. Most physical laws are computable, and thus algorithmic. However, Penrose determined that wave function collapse was a prime candidate for a non-computable process. In quantum mechanics, particles are treated differently from the objects of classical mechanics. Particles are described by wave functions that evolve according to the Schrödinger equation. Non-stationary wave functions are linear combinations of the eigenstates of the system, a phenomenon described by the superposition principle. When a quantum system interacts with a classical system—i.e. when an observable is measured—the system appears to collapse to a random eigenstate of that observable from a classical vantage point.

If collapse is truly random, then no process or algorithm can deterministically predict its outcome. This provided Penrose with a candidate for the physical basis of the non-computable process that he hypothesized to exist in the brain. However, he disliked the random nature of environmentally induced collapse, as randomness was not a promising basis for mathematical understanding. Penrose proposed that isolated systems may still undergo a new form of wave function collapse, which he called objective reduction (OR).[7]

> Penrose sought to reconcile general relativity and quantum theory using his own ideas about the possible structure of spacetime.[18][*page needed*][19] He suggested that at the Planck scale curved spacetime is not continuous, but discrete. He further postulated that each separated quantum superposition has its own piece of spacetime curvature, a blister in spacetime. Penrose suggests that gravity exerts a force on these spacetime blisters, which become unstable above the
>
> Planck scale of and collapse to just one of the possible states. The rough threshold for OR is given by Penrose's indeterminacy principle, where:

- is the time until OR occurs,
- is the gravitational self-energy or the degree of spacetime separation given by the superpositioned mass, and
- is the reduced Planck constant.

Thus, the greater the mass–energy of the object, the faster it will undergo OR and vice versa. Mesoscopic objects could collapse on a timescale relevant to neural processing.[7][*additional citation(s) needed*]

An essential feature of Penrose's theory is that the choice of states when objective reduction occurs is selected neither randomly (as are choices following wave function collapse) nor algorithmically. Rather, states are selected by a "non-computable" influence embedded in the Planck scale of spacetime geometry. Penrose claimed that such information is Platonic, representing pure mathematical truths, which relates to Penrose's ideas concerning the three worlds: the physical, the mental, and the Platonic mathematical world. In *Shadows of the Mind* (1994), Penrose briefly indicates that this Platonic world could also include aesthetic and ethical values, but he does not commit to this further hypothesis.[19]

The Penrose–Lucas argument was criticized by mathematicians,[20][21][22] computer scientists,[12] and philosophers,[23][24][9][10][11] and the consensus among experts in these fields is that the argument fails,[25][26][27] with different authors attacking different aspects of the argument.[27][28] Minsky argued that because humans can believe false ideas to be true, human mathematical understanding need not be consistent and consciousness may easily have a deterministic basis.[29] Feferman argued that mathematicians do not progress by mechanistic search through proofs, but by trial-and-error reasoning, insight and inspiration, and that machines do not share this approach with humans.[21]

Orch OR

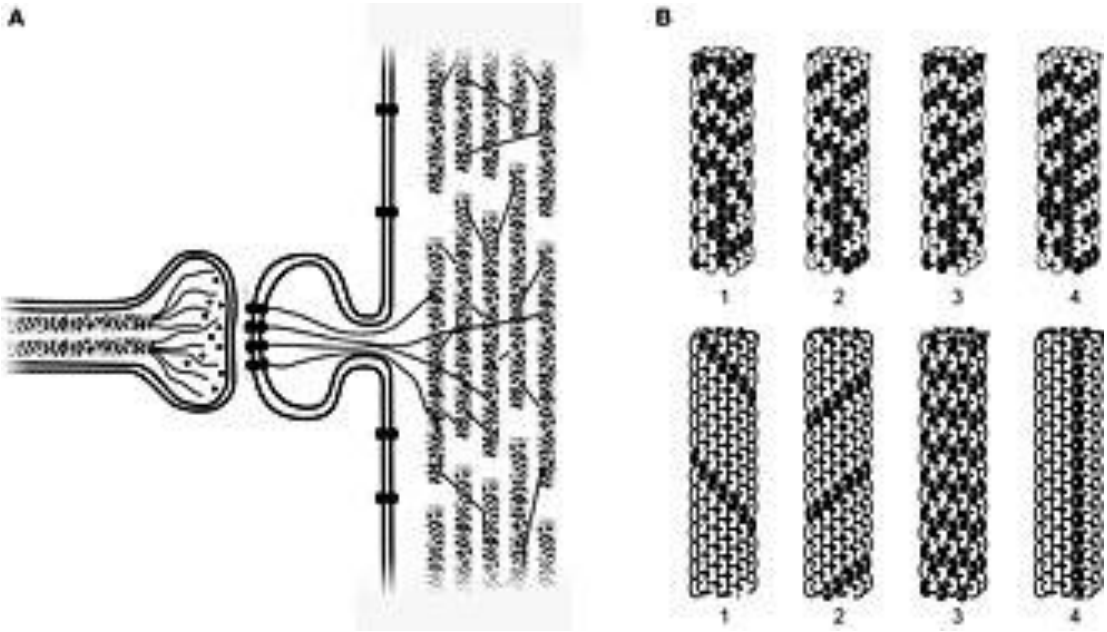
Penrose outlined a predecessor to Orch OR in *The Emperor's New Mind*, coming to the problem from a mathematical viewpoint and in particular Gödel's theorem, but lacked a detailed proposal for how quantum processes could be implemented in the brain. Stuart Hameroff separately worked in cancer research and anesthesia, which gave him an interest in brain processes. Hameroff read Penrose's book and suggested to him that microtubules within neurons were suitable candidate sites for quantum processing, and ultimately for consciousness.[30][31] Throughout the 1990s, the two collaborated on the Orch OR theory, which Penrose published in *Shadows of the Mind* (1994).[19]

Hameroff's contribution to the theory derived from his study of the neural cytoskeleton, and particularly on microtubules.[31] As neuroscience has progressed, the role of the cytoskeleton and microtubules has assumed greater

importance. In addition to providing structural support, microtubule functions include axoplasmic transport and control of the cell's movement, growth and shape.[31]

Orch OR combines the Penrose–Lucas argument with Hameroff's hypothesis on quantum processing in microtubules. It proposes that when condensates in the brain undergo an objective wave function reduction, their collapse connects noncomputational decision-making to experiences embedded in spacetime's fundamental geometry. The theory further proposes that the microtubules both influence and are influenced by the conventional activity at the synapses between neurons.

Microtubule computation



A: An axon terminal releases neurotransmitters through a synapse and are received by microtubules in a neuron's dendritic spine.
B: Simulated microtubule tubulins switch states.[1]

Hameroff proposed that microtubules were suitable candidates for quantum processing.[31] Microtubules are made up of tubulin protein subunits. The tubulin protein dimers of the microtubules have hydrophobic pockets that may contain delocalized π electrons. Tubulin has other, smaller non-polar regions, for example 8 tryptophans per tubulin, which contain π electron-rich indole rings distributed throughout tubulin with separations of roughly 2 nm. Hameroff claims that this is close enough for the tubulin π electrons to become quantum entangled.[32] During entanglement, particle states become inseparably correlated.

Hameroff originally suggested in the fringe *Journal of Cosmology* that the tubulin-subunit electrons would form a Bose–Einstein condensate.[33] He then proposed a Frohlich condensate, a hypothetical coherent oscillation of dipolar

molecules. However, this too was rejected by Reimers's group.[34] Hameroff and Penrose contested the conclusion, considering that Reimers's microtubule model was oversimplified.[35]

Hameroff then proposed that condensates in microtubules in one neuron can link with microtubule condensates in other neurons and glial cells via the gap junctions of electrical synapses.[36][37] Hameroff proposed that the gap between the cells is sufficiently small that quantum objects can tunnel across it, allowing them to extend across a large area of the brain. He further postulated that the action of this large-scale quantum activity is the source of 40 Hz gamma waves, building upon the much less controversial theory that gap junctions are related to the gamma oscillation.[38]

Related experimental results

Superradiance

In a study Hameroff was part of, Jack Tuszyński of the University of Alberta demonstrated that anesthetics hasten the duration of a process called delayed luminescence, in which microtubules and tubulins re-emit trapped light. Tuszyński suspects that the phenomenon has a quantum origin, with superradiance being investigated as one possibility (in a 2024 study, superradiance was confirmed to occur in networks of tryptophans, which are found in microtubules).[39][40] Tuszyński told *New Scientist* that "We're not at the level of interpreting this physiologically, saying 'Yeah, this is where consciousness begins,' but it may."[41]

The 2024 study, called Ultraviolet Superradiance from Mega-Networks of Tryptophan in Biological Architectures and published in *The Journal of Physical Chemistry*, confirmed superradiance in networks of tryptophans.[39][40] Large networks of tryptophans are a warm and noisy environment, in which quantum effects typically are not expected to take place.[39] The results of the study were theoretically predicted and then experimentally confirmed by the researchers.[39][40] Majed Chergui, who led the experimental team, stated that "It's a beautiful result. It took very precise and careful application of standard protein spectroscopy methods, but guided by the theoretical predictions of our collaborators, we were able to confirm a stunning signature of superradiance in a micron-scale biological system."[39] Marlan Scully, a physicist well-known for his work in the field of theoretical quantum optics, said "We will certainly be examining closely the implications for quantum effects in living systems for years to come."[39] The study states that "by analyzing the coupling with the electromagnetic field of mega-networks of tryptophans present in these biologically relevant architectures, we find the emergence of collective quantum optical effects, namely, superradiant and subradiant eigenmodes. ... our work demonstrates that collective and cooperative UV excitations in mega-networks of tryptophans support robust quantum states in protein aggregates, with observed consequences even under thermal equilibrium conditions."[40]

Microtubule quantum vibration theory of anesthetic action

In an experiment, Gregory D. Scholes and Aarat Kalra of Princeton University used lasers to excite molecules within tubulins, causing a prolonged excitation to diffuse through microtubules farther than expected, which did not occur when repeated under anesthesia.[42] However, diffusion results have to be interpreted carefully, since even classical diffusion can be very complex due to the wide range of length scales in the fluid filled extracellular space.[43]

At high concentrations (~5 MAC) the anesthetic gas halothane causes reversible depolymerization of microtubules.[44] This cannot be the mechanism of anesthetic action, however, because human anesthesia is performed at 1 MAC. (Neither Penrose or Hameroff claim that depolymerization is the mechanism of action for Orch OR.)[45][46] At ~1 MAC halothane, reported minor changes in tubulin protein expression (~1.3-fold) in primary cortical neurons after exposure to halothane and isoflurane are not evidence that tubulin directly interacts with general anesthetics, but rather shows that the proteins controlling tubulin production are possible anesthetic targets.[47] Further proteomic study reports 0.5 mM [^{14}C]halothane binding to tubulin monomers alongside three dozens of other proteins.[48] In addition, modulation of microtubule stability has been reported during anthracene general anesthesia of tadpoles.[49] The study called Direct Modulation of Microtubule Stability Contributes to Anthracene General Anesthesia claims to provide "strong evidence that destabilization of neuronal microtubules provides a path to achieving general anesthesia".[49]

A highly disputed theory put forth in the mid-1990s by Hameroff and Penrose posits that consciousness is based on quantum vibrations in tubulin/microtubules inside brain neurons. Computer modeling of tubulin's atomic structure[50] found that anesthetic gas molecules bind adjacent to amino acid aromatic rings of non-polar π-electrons and that collective quantum dipole oscillations among all π-electron resonance rings in each tubulin showed a spectrum with a common mode peak at 613 THz.[51] Simulated presence of 8 different anesthetic gases abolished the 613 THz peak, whereas the presence of 2 different nonanesthetic gases did not affect the 613 THz peak, from which it was speculated that this 613 THz peak in microtubules could be related to consciousness and anesthetic action.[51]

Another study that Hameroff was a part of claims to show "anesthetic molecules can impair π-resonance energy transfer and exciton hopping in 'quantum channels' of tryptophan rings in tubulin, and thus account for selective action of anesthetics on consciousness and memory".[52]

In a study published in August 2024, an undergraduate group led by a Wellesley College professor found that rats given epothilone B, a drug that binds to microtubules, took over a minute longer to fall unconscious when exposed to an anesthetic gas.[53]

Criticism

Orch OR has been criticized both by physicists[14][54][34][55][56] and neuroscientists[57][58][59] who consider it to be a poor model of brain physiology. Orch OR has also been criticized for lacking explanatory power; the philosopher Patricia Churchland wrote, "Pixie dust in the synapses is about as explanatorily powerful as quantum coherence in the microtubules."[60]

David Chalmers argues against quantum consciousness. He instead discusses how quantum mechanics may relate to dualistic consciousness.[61] Chalmers is skeptical that any new physics can resolve the hard problem of consciousness.[62][63][64] He argues that quantum theories of consciousness suffer from the same weakness as more conventional theories. Just as he argues that there is no particular reason why particular macroscopic physical features in the brain should give rise to consciousness, he also thinks that there is no particular reason why a particular quantum feature, such as the EM field in the brain, should give rise to consciousness either.[64]

Endogenous ferritin quenches microtubule radiance, which may prevent generation of UV biophotons

While some of the studies noted above purport to show superradiance and an influence of anesthetics on decreasing excitation diffusion through microtubules, those studies were performed under artificial conditions that failed to include microtubule associated proteins like ferritin,[65] which quenches microtubule superradiance ("Intrinsic Trp fluorescence of tubulin is quenched upon binding of Apf with tubulin and indicates that its binding with tubulin perturbs tubulin conformation in the vicinity of Trp residues").[66] Extensive evidence that was published prior to those studies establishes that ferritin interacts with microtubules in vivo and is essential for microtubule stability and function[67] and should have been addressed by those studies.

For example, some of the evidence that those studies failed to consider includes:

- Studies of biophotons in the human body fail to find any evidence of UV biophotons.[68] In contrast, at least one of the studies cited above that is relied on as evidence of microtubule superradiance in support of Orch-OR relies on earlier studies of UV biophotons measured in single-celled organisms like *E. coli* and respiratory-deficient yeast as the basis for its contention that such biophotons are present in cells.[69][70] That study also used UV-vis equipment with a light source that can generate 10^{20} photons per second. It is difficult to understand how it can be argued that tests that are conducted under artificial conditions with light sources that generate 10^{20} photons per second are relevant to neurons, where there is no evidence that UV

biophotons are even generated. Those tests are not even relevant to *E. coli* or yeast.

- Ferritin in the human body absorbs UV from external sources at least in the skin and in the cornea, where the levels of UV photons are much higher than measured biophoton levels of UV, even in *E. coli* and yeast.[71][72] Endogenous ferritin in neurons would absorb UV biophotons that might be emitted from chemical processes (at levels that are too low to measure), and those UV biophotons would not even reach microtubules to cause superradiance or energy transport.
- Ferritin in animals is associated with the generation of biophotons above the UV range, which is further evidence that ferritin in cells might absorb UV biophotons or otherwise prevent them from being generated in animals.[73]
- Ferritin is needed for microtubule stability, and microtubules fall apart when ferritin is not present. Ferritin cannot be ignored as part of any serious study that purports to demonstrate that microtubule superradiance or energy transport has any place in biology, because ferritin is present in almost every cell of every plant and animal.[74]
- Ferritin contains tryptophan residues, the same material in microtubules that is supposed to cause microtubule superradiance.[75] According to one of the studies cited above, microtubule superradiance is based on special configurations of tryptophan residues. The failure of that study to consider additional ferritin tryptophan residues in the vicinity of microtubule tryptophan residues means that the study is not relevant to cellular environments that include ferritin (which is basically every cell). As noted above, ferritin perturbs tubulin in the vicinity of tryptophan residues, which invalidates an *a priori* assumption of that study.
- Ferritin has stronger ionic interaction with microtubules than the anesthetics that were used in one of the studies cited above and has electrical and magnetic properties that those anesthetics lack.[76][77][78] Even if anesthetics interact with microtubules, ferritin has stronger interactions with microtubules, which may explain why ferritin is able to quench microtubule fluoresence.

In summary, no serious attempt appears to have been made to investigate whether Orch-OR might be present in neurons by the tests that are relied as evidence of Orch-OR. Rather, the tests appear designed to prop up the extraordinary claims of Orch-OR upon thin reeds of evidence, using unrealistic levels of UV light, artificial environments and by excluding important cellular substances that would prevent microtubule superradiance and energy transport.

Decoherence in living organisms

In 2000 Max Tegmark claimed that any quantum coherent system in the brain would undergo effective wave function collapse due to environmental interaction long before it could influence neural processes (the "warm, wet and noisy"

argument, as it later came to be known).[14] He determined the decoherence timescale of microtubule entanglement at brain temperatures to be on the order of femtoseconds, far too brief for neural processing. Christof Koch and Klaus Hepp also agreed that quantum coherence does not play, or does not need to play any major role in neurophysiology.[15][16] Koch and Hepp concluded that "The empirical demonstration of slowly decoherent and controllable quantum bits in neurons connected by electrical or chemical synapses, or the discovery of an efficient quantum algorithm for computations performed by the brain, would do much to bring these speculations from the 'far-out' to the mere 'very unlikely'."[15]

In response to Tegmark's claims, Hagan, Tuszynski and Hameroff claimed that Tegmark did not address the Orch OR model, but instead a model of his own construction. This involved superpositions of quanta separated by 24 nm rather than the much smaller separations stipulated for Orch OR. As a result, Hameroff's group claimed a decoherence time seven orders of magnitude greater than Tegmark's, although still far below 25 ms. Hameroff's group also suggested that the Debye layer of counterions could screen thermal fluctuations, and that the surrounding actin gel might enhance the ordering of water, further screening noise. They also suggested that incoherent metabolic energy could further order water, and finally that the configuration of the microtubule lattice might be suitable for quantum error correction, a means of resisting quantum decoherence.[79][80]

In 2009, Reimers *et al.* and McKemmish *et al.,* published critical assessments. Earlier versions of the theory had required tubulin electrons to form either Bose–Einstein or Frohlich condensates, and the Reimers group noted the lack of empirical evidence that such could occur. Additionally, they calculated that microtubules could only support weak 8 MHz coherence. McKemmish *et al.* argued that aromatic molecules cannot switch states because they are delocalised, and that changes in tubulin protein conformation driven by GTP conversion would result in a prohibitive energy requirement.[54][34][55]

In 2022, a group of Italian physicists conducted several experiments that failed to provide evidence in support of a gravity-related quantum collapse model of consciousness, weakening the possibility of a quantum explanation for consciousness.[81][82]

Biology-based criticisms have been offered, including a lack of explanation for the probabilistic release of neurotransmitter from presynaptic axon terminals[83][84][85] and an error in the calculated number of tubulin dimers per cortical neuron.[86]In 2014, Penrose and Hameroff published responses to some criticisms and revisions to many of the theory's peripheral assumptions, while retaining the core hypothesis.[2][6]