Biol Philos (2008) 23:157–177
DOI 10.1007/s10539-007-9079-5

# Nano-intentionality: a defense of intrinsic intentionality

**W. Tecumseh Fitch**

Received: 23 January 2007 / Accepted: 18 June 2007 / Published online: 25 August 2007


**Abstract** I suggest that most discussions of intentional systems have overlooked an important aspect of living organisms: the intrinsic goal-directedness inherent in the behaviour of living eukaryotic cells. This goal directedness is nicely displayed by a normal cell's ability to rearrange its own local material structure in response to damage, nutrient distribution or other aspects of its individual experience. While at a vastly simpler level than intentionality at the human cognitive level, I propose that this basic capacity of living things provides a necessary building block for cognition and high-order intentionality, because the neurons that make up vertebrate brains, like most cells in our body, embody such capacities. I provisionally dub the capacities in question "nano-intentionality": a microscopic form of "aboutness". The form of intrinsic intentionality I propose is thoroughly materialistic, fully compatible with known biological facts, and derived non-mysteriously through evolution. Crucially, these capacities are not shared by any existing computers or computer components, and thus provide a clear, empirically-based distinction between brains and currently existing artificial information processing systems. I suggest that an appreciation of this aspect of living matter provides a potential route out of what may otherwise appear to be a hopeless philosophical quagmire confronting information-processing models of the mind.

**Keywords** Evolution of mind · Evolution of cognition · Intentionality · Intrinsic intentionality

## Introduction

A core problem for the scientific study of mind comes down to this: how is it possible for an arrangement of matter (e.g. the brain) to embody a mind that has subjectively felt

W. T. Fitch (✉)
School of Psychology, St Mary's Quad, University of St Andrews, Rm 2.57, St Andrews, Fife KY16 9JP, UK
e-mail: wtsf@st-andrews.ac.uk

experiential states (thoughts, pains, desires, etc)? The problem is not with understanding arrangements of matter that do complex things (a Cuisinart or a laptop or a heart do that, without residual mystery) or even arrangements of matter that process information (a computer or a brain non-mysteriously does that). The problem boils down to the question of how something objective like *my brain*, that one could weigh, slice and analyze in microanatomical detail, can produce something subjective like *me* (that—I assure you—has thoughts, pains, pleasures, beliefs, desires—just as I assume *you* do).

Philosophers of mind have traditionally (since Brentano) couched this problem in terms of "intentionality" (a technical term with only an indirect relationship to its ordinary English meaning of willful, conscious, action). Intentionality in the philosopher's sense adopted here refers to the "aboutness" of certain types of things in the world, including both minds, and of artifacts such as words or other meaningful linguistic units. Daniel Dennett has cast the difficulty in sharp relief using a distinction between "derived intentionality" and "intrinsic intentionality" (Dennett 1987). Canonically, *derived intentionality* is that form of aboutness possessed by functional objects (especially words, but also thermostats, Cusinarts, etc.) which are "about" something by virtue of their *intended* function (p 29 ff, Dennett 1987). Whatever aboutness they have is derivative of the "real" intentionality of their mindful creators. Biological "artefacts" (e.g. hearts or kidneys) possess a similar aboutness—a function—but in this case their function is not that foreseen in the mind of their creator, but rather by the foresightless "blind watchmaker" of natural selection (Millikan 1987). By this analogy, hearts or kidneys also have "derived" intentionality. Dennett's radical suggestion is to extend this notion of evolutionarily-derived intentionality to ourselves as well: to accept the fact that we are products of natural selection and thus that our aboutness, too, is simply derived via evolution.

Dennett contrasts this pillar of his philosophy of mind with the viewpoint of many other philosophers of mind (Fodor, Searle, Dretske, Kripke, etc) concerning the existence of a different type of intentionality (Dennett 1987). Dennett quotes Fodor as saying "after all Searle is right about one thing: artifacts don't have original or intrinsic intentionality: only derived intentionality" (Dennett 1987). Such philosophers argue that we have something—*intrinsic intentionality*—that words or machines don't. The thrust of Dennett's argument is that this "something" (which he terms variously "original" and "intrinsic" intentionality) is an illusion, a deep philosophical mistake that derives from our unwillingness to fully bite the bullet and accept natural selection (and its products, especially ourselves) for the blind and goal-less process that it is. Dennett criticizes intrinsic intentionality with arguments that need to be taken seriously, arguments which seem to force us into a corner where we must choose between the following propositions: either accept that *all* intentionality, including our own, is derivative (and then admit that a thermostat has a little bit of intentionality, too), or we retreat from this unpalatable option into a mysterian, outmoded belief in "original/intrinsic intentionality", a concept that under close scrutiny leads to contradiction and paradox at every turn.

In this paper I will argue against Dennett's conclusion by focusing on a causal power related to "intrinsic intentionality" that seems to have been overlooked in previous discussions. The key to my argument is the recognition of a specific capacity characterizing eukaryotic cells (the type of highly-organized cells with DNA sequestered in a nucleus, organelles, a cytoskeleton, etc, that make up a huge variety of life forms including amoebae, mushrooms, redwoods, and humans). I provisionally dub this causal power "nano-intentionality"—a microscopic form of aboutness, inherent in individual eukaryotic cells, that includes a goal-directed capacity to respond adaptively to novel circumstances. The core causal power underlying nano-intentionality is the cell's ability to arrange and rearrange its own molecules in a locally-functional manner, thus preserving and extending its individual existence, depending on local and perhaps somewhat novel circumstances. Crucially, this capacity is as characteristic of a neuron in the brain as it is of a free-living amoeba. Both deal, as semi-autonomous individual cells, with their local circumstances, and when novelty is coped with successfully, the cell can "remember" a solution by changing its own physical structure. This specific capacity is an important characteristic of virtually any eukaryotic cell and is related to, but much more specific than, Aristotelian "telos" or Schopenhauerian "will". I will argue here that humans possess intrinsic intentionality, and that it necessarily builds upon the nano-intentionality of the cells of our brains. They are not the same thing, however: the relationship between nano- and intrinsic intentionality is quite complex and indirect. By unpacking the linkage between nano-intentionality and mental activity I suggest that we can avoid both horns of Dennett's dilemma for the low, low price of taking seriously some uncontroversial biological facts about eukaryotic cells and the vertebrate brain.

As a biologist I accept without argument the "naturalist" position—as argued by Dennett, (Millikan 1987) and many others—that intentionality of all sorts is ultimately the result of evolution via natural selection. I also assume a non-vitalist stance (as have most biologists since Humboldt): a cell is "just a machine". There is no non-physical *élan vital* that sets off living things from other complex physical objects, and all of a cell's functions can, in principle, be derived from the chemistry and physics of its components, and thus "reduced" to physics in the same way as the functions of carburetors and laptops. Although our understanding of this reduction remains incomplete, the immense and rapid progress of cellular and molecular biology in the last half-century leave little grounds for doubt of this principle. However, the "cell = machine" truism obscures an equally important fact: that a eukaryotic cell is unlike any machine ever produced by humans. A cell has specific, causal powers, possessed by no currently available machine, and it is these powers I wish to bring into focus with the term "nano-intentionality". Eukaryotic cells respond adaptively and independently to their environment, rearranging their molecules to suit their local conditions, based on past (individual and species) history. A transistor or a thermostat does not—nor do the most complex machines currently available. This is a *practical* difference between cells and machines (I can see no reason *in principle* that machines must lack such causal powers), but it is nonetheless a profound one. I suggest here that it is by virtue of this difference, this possession of nano-intentionality by the cellular machines that make up our bodies and brains, that we, as whole individuals, possess "intrinsic" intentionality—and a laptop does not. The purpose of the following paper is to show why, in principle, this is a possibility that should be considered seriously by philosophers of mind and neuroscientists.

## Nano-intentionality in context

### Situating the question

How can matter have, or make, mind? The oldest answers to this question assume that mind and matter are two separate things, connected only coincidentally in living humans: "we have a mind/spirit because the gods gave us one, and it will live on when our bodies are gone". Such mythic/religious answers simply replace one mystery with another, but early attempts to be scientific about the problem, such as those of Descartes, inherited their language and intellectual baggage. Descartes' solution, consistent with the science of his time, was that MOST matter (including everything that happens in any animal besides man) is nothing more than a machine—despite appearances non-humans do not in fact have subjective pains or beliefs. The fact that a dog will yelp and seek to avoid precisely those stimuli that cause objective tissue damage and pain when applied to humans was not very convincingly swept under the carpet, and the critical “additional” element that makes it *pain* for us and just *tissue damage* for the dog was held to be just old-fashioned *esprit* (mind, spirit or soul), given to us by God. Given its persistently mysterious character, it is more than a bit surprising that Descartes’ “solution” has had such lasting influence, remaining a starting point for discussions even today. Historically, by ceding the study of matter and most of the biological world to the natural sciences, Descartes did biology a huge but temporary favour. But it was only a matter of time before reductionist science, having conquered the heart and liver (and now the genetic code and the biophysics of neuronal firing) would return to the problem of the brain/mind connection that Descartes only postponed without answering (he did have an answer involving the pineal gland, but it is hard to imagine that even Descartes took this answer very seriously, and no one ever has since).

By the mid-20th century, human technology advanced to the state where engineers could build a machine that proved some basic theorems or played a decent game of chess, and the mind/body problem once more reared its fascinating head (spotted first in its modern guise by Alan Turing). In the same way that the function of the heart is pumping, the function of the brain is surely thinking. But playing chess is thinking, is it not? So, if mind and brain are related just like pumping and hearts, doesn’t a chess-playing computer *already* have a mind (if only a rather single-minded one)? The intuition of your average person on the street is that this suggestion is manifestly ridiculous, and that of course a computer doesn’t have a mind, even one that beats Gary Kasparov. Nonetheless, many cognitive scientists today would answer that, yes, the computer does have *a little bit* of mind, and many (perhaps most) philosophers would agree. They might feel uncomfortable about it, but many scientists who study the mind basically feel forced to this conclusion by their most dearly-held assumptions about materialism, scientific objectivity, and logic. I want to rethink, and ultimately reject, this conclusion.

## Situating the answer

I argue that people (and other vertebrates, at least) have a subjective, first-person mind, and that this is a real thing, independent of anyone else’s opinion about the matter. I think (feel, believe, desire...), and therefore I am (a subjective, first person entity), and I do so irrespective of what anyone else thinks or believes about me. Unlike Descartes, I argue that the same is true of a dog—but not of a laptop. This intentionality is intrinsic to brains, and not dependent on me or the dog or anyone else adopting an “intentional stance” towards them (contra Dennett 1987). Thus I argue that this subjective side of consciousness is actually a real thing in the world, not to be explained away as an illusion or misconception or deep philosophical mistake. Therefore, I am arguing for a version of what Dennett has castigated as “intrinsic intentionality”. As Dennett concedes, I’ve got time-honoured intuition on my side, along with some prominent philosophers of mind like John Searle who have made the case clearly with clever parables like the Chinese Room (Searle 1980). We all know, nevertheless, that intuition is no infallible guide to truth. This much has already been clearly said.

However, a biological point has been submerged in much of the philosophical debate,[1] concerning some specific functions typifying biological systems at the cellular level (and thus incorporated into organs like the brain that are composed of such cells). I know of no normal word for this type of function, so I will adapt the philosopher's term "intentionality" (the "aboutness" of certain things, like words or thoughts) to refer to it. Part of being a word is being "about" something (e.g. the string of phonemes "cat" is prototypically "about" small domesticated carnivores that meow), and part of being a *thought* is also being "about" something in a similar way. Thus (presumably) part of being a pattern of neural firing that *constitutes* a thought is being about something too. But if so, isn't the pattern of electron movement in a silicon chip that is executing a chess-playing program "about" something, in precisely the same way?

My answer to this question is no. Although, adopting Dennett's intentional stance, we can certainly say that the computer's flux is "about" playing chess, because that's what the computer programmer responsible for the program thinks, we could equally say that the flux is about assuring the programmer's career, or about making Bill Gates more money, or many other things. But as far as the computer *itself* is concerned, it is not intrinsically "about" anything, because *all* of its meaningful aboutness is derived or "borrowed" from the programmer, or imputed to it by a user. As far as the computer itself knows or cares, the electron flux could just as well be about solving a differential equation or compressing the digital images from my summer vacation. The chip does the same thing regardless, and in fact ALWAYS does this same one thing, unless it is broken in which case it simply generates heat. I am suggesting, in contrast, that the ion flux in the human nervous system HAS a form of intrinsic intentionality, by virtue of a type of "aboutness" of the cells that generate and respond to these ions.

I will call this specific capacity "nano-intentionality". Without nano-intentionality, I argue, intentionality proper can never emerge. Without such a capacity, all of the information processing in the world will not make a system intentional. Indeed, the belief that it should, I think, is a trap laid by the functionalist assumption of the implementation-independence of mind that needs reexamination. The brain is about the body, and has causal powers over the body, in a very specific way that builds upon the way in which cellular behaviour is about the "body" or physical form of the cell. So, rather than intentionality being something that accrues gradually over a series of evolutionary stages (Dennett 1996), I am suggesting that a certain specific type of physical causal power needs to be present, right from the beginning, for intentionality to exist at all (and thus conscious thought, meaningful language, and the other things that build upon and flow from it). My argument hinges on some specific aspects of cellular behaviour as crucial intervening explanatory factors in understanding how brains make minds. In this way it directly parallels recent insights in evolutionary developmental biology about the crucial intervening role of cellular behaviour in development and evolutionary change (Carroll et al. 2005; Kirschner and Gerhart 2005).

[1] However, I note that Arnold Schopenhauer's core point about what he termed the "will" presages many of the points I make here Schopenhauer (1819). Schopenhauer correctly saw the error of Kant's key claim that we are wholly and irrevocably separated from the actual physical reality of things—Kant's *Ding in sich*. According to Schopenhauer, our minds *do* in fact make direct contact with the actual reality of lifeless physical matter—in the single specific domain of our own bodies. Schopenhauer's convincing if underappreciated argument against Kant provides a philosophical precursor of the biologically-based argument I present here. Unfortunately, Schopenhauer goes one step too far I think by ascribing this "will" to non-living matter as well; by my understanding "will" in his sense is the crucial quality distinguishing the living from the non-living, and "will" is closely allied to what I term nano-intentionality.

## Machines, cells and nano-intentionality

A crucial difference between a cell (including but not limited to a neuron) and a transistor on a silicon chip is that the former arrangement of matter can autonomously and adaptively modify itself in response to its circumstances, whereas the latter cannot. An everyday example of this biological capacity is provided by the healing response: a damaged organism can often stem the loss of precious bodily fluids, stitch itself up, and (with some scar perhaps) continue living. We all witness this capacity regularly in our own bodies, and it is worth stopping for a moment to realize how amazing it is. We know there can be no "blueprint" of the body built into the genes (no way that $10^9$ base pairs can detail the micro-plan of a body containing $10^{15}$ cells), so how do these repairing cells "know" what to do, and when and where to do it? Many organisms (take salamanders for example) can regrow entire body parts like lost limbs. And note that this capacity is not only true of complex multicellular organisms like trees or vertebrates: a single-celled foraminiferan can do the same at the microscopic level (Tartar 1961; Grell 1973; Goldstein 1999).

In the same way that a plant adjusts its form to local lighting conditions, a neuron continually adjusts itself to its local, individual circumstances in the brain—producing more neurotransmitter when it runs low, extending spines out to make more effective contact with a preferred neighbor and withdrawing dendritic branches from noisy uncorrelated neighbors, adjusting its pattern of firing to the incoming flow of neurotrophins, and even curling up and neatly committing suicide when unable to integrate itself into its local processing environment. This dynamic rearrangement of its own matter is something that all neurons do, and indeed virtually all cells in our body do (omitting, e.g. red blood cells which lack their own nucleus and are more a bag of hemoglobin than a living thing). Not only do cells modify themselves, but they do so adaptively (in the physiological sense)—they autonomously arrange their form in such a way as to optimize their ability to perform certain quite specific functions. When some abstract goal remains unsatisfied, they try new behaviours, mostly at random, and "remember" the one(s) that work(s). Furthermore, such local adaptive changes can be incorporated into the cell's form for deployment in future similar circumstances. The cell can "remember" its successes and thus in this limited way can "learn" from its individual experience. This capacity to respond to circumstances includes *novel* circumstances (in a constrained range): circumstances that were perhaps never encountered in the lifetime of the species, or at least never led to inherited changes in the cell's DNA. They are in this sense "unforeseen" by natural selection, although of course the capacity to implement this limited constrained novelty-handling has evolved—just as the ability to learn has evolved, at a much higher level of organization, but the specific knowledge learned has not.

Contrast with this contemporary machinery of any sort. While it is rather easy to make a machine that modifies itself (e.g. if I rewire my oven, I can probably make it melt itself into a lump of iron), it's much harder to make a machine that changes itself *for the better*. But if we did succeed in this (and I think that would be possible with tomorrow's technology) we would have made a machine that shares with cells one of their most basic characteristics: the ability to autonomously, adaptively rearrange their microstructure to suit some function. We have many complex machines (cars, Boeing 747s, satellites) that can sense damage to their components, and some even make preprogrammed corrective response. Combine that with a small, mobile welding/rewiring robot, and an AI inference engine capable of dealing with unforeseen damage types, and we could build a self-healing machine. Such a self-modifying capacity is built-in to virtually every living eukaryotic cell on the planet: and it is built-in to every one of the $10^{11}$ neurons in our brains (along with the glia and other supporting cells).

But you may be thinking that surely *any* complex machine rearranges its form to suit its function. The governor on a steam engine, the distributor on my old Honda or even a simple thermostat changes its physical configuration in order to accomplish its function. What's the difference from a cell? One difference is that these machines consist of a set of parts of fixed shape that can change their positions in relation to one another, but cannot (adaptively) change their own form. The molecular configuration (shape) of a gear or shaft or thermostat coil's is impressed upon it at the time of manufacture, and from that point on, degradation or random accretion is all that can occur. Any repairs or modifications must come from outside the system. All the gears in my car's transmission are able to do *by themselves* is grind themselves down—somone else has to be hired at ridiculous expense to repair or replace them. A computer chip (or any other solid-state circuit) can't even change its configuration of parts: it is truly a static material object, and all of its relevant dynamics are in the flow of electrons through it. You can think of an Intel chip as an elaborate fixed system of pipes, like a vast water supply system where different pressures make different flows happen at different times, but the physical form remains absolutely static, and even small changes in its form will lead to failure. Another difference is in the type of adjustments available: while a thermostat operates on a pre-defined control axis (temperature) with preset "goals" (the set point) and inputs (the thermometer), most living things can actually modify their input–output relations based on information about the current situation, update their actions based on these factors, and thus engage successfully with problems for which they were not specifically "designed" by evolution.[2]

In contrast to a silicon chip, your brain is constantly changing at a microscopic, molecular level, not just in the trivial sense of blood or nutrients flowing around in it but in the deeply important sense that its very purpose in life—processing information—entails that each component neuron rearrange its axons and dendrites, and change the form of its synapses on a moment-to-moment basis. If you remember anything about reading this essay, or eating today's dinner, tomorrow, it is because the neurons composing your brain actually reconfigured their material structure. This, I argue, is the crucial difference between living machines and all the man-made machines currently available. That it is *a* difference is a matter of fact and hardly arguable. However, the suggestion that it is *the* difference that will help resolve the issue of intentionality or help "solve the mind/body problem" will probably seem far-fetched. This is where things get interesting.

## A cascade of causal powers: a reductionist view of intentionality

Dennett has suggested that, although a thermostat is not intentional in a human sense, it has a stripped-down form of aboutness, no different fundamentally from our own but differing only in complexity (p. 22 ff. Dennett 1987). I believe that this argument is incorrect for mechanical switches, but I accept much of Dennett's logic otherwise. I will suggest that

[2] I thank Phillip Pettit for pointing out this key difference between the flow-diagram for intentional systems as normally conceived, and the intentionality of living things. A nice analogy is the behavior of a fish or frog captured in a trap—they will adopt the (implicit) goal of "escape", generate mostly random motoric behaviour to achieve this goal, and repeat until they are free. But "escape" and "freedom" are not goals of evolution or "free-floating rationales" built in by past phylogenetic history. They are rough-and-ready responses to the novel and undesireable situation the individual finds itself in. Pettit suggests the nice term "intra-active intentionality" for this more fluid conception of intentionality.

neurons *do* have a basic form of intentionality, but in the same way that one might say that a single gas molecule in an otherwise empty chamber has "just a bit" of pressure. However, just as we need a large number of molecules before any of the properties we normally attribute to "gases" as physical entities can be sensed or measured, we need many cells interacting in specific ways to get minds. Using the same term (intentionality/ "aboutness") to refer to the cellular-level phenomenon and the mental phenomenon may be slightly misleading—like saying that the hardness of a diamond inheres in the strength of the bond between two carbon atoms, or that the wetness of water inheres in the weakness of covalent bonds. These low-level physical properties of bonds as constitutive of the mass-level properties of hardness or wetness. These classic examples of "reduction" in the physicist's sense serve as micro-level stand-ins that *en masse* constitute the macro-level property that can be sensed or measured.

The core hypothesis of this essay is that full mental intentional capacities are undergirded by the nano-intentional aboutness possessed by cells, although cells are obviously not themselves mental. When combined properly into large interconnected systems, this combined mass-action of cellular nano-intentionality yields intrinsic intentionality in the typical philosopher's sense, as well as both consciousness and the efficacy of our subjectively sensed self to move the body and perform other acts of will ("intentionality" in its general English sense), just as the mass and velocity of gas molecules contained in a volume are constitutive of its temperature, pressure and weight. I argue that nano-intentionality of neurons stands in a similar sort of causal relation to the subjective sense of "aboutness" we all have, the first person awareness of sensations and beliefs. However, the link between cellular nano-intentionality and mind is far less direct than that between momentum and temperature, requiring several intervening levels of causal organization (as detailed below). I differ from Dennett, I think, only in granting this constitutive capacity to a neuron or to an amoeba, but not to a transistor or to a thermostat. I have already said why, practically speaking: because cells are capable of rearranging their molecules based on individual circumstances in an autonomous and adaptive fashion, while the latter are not. In the remainder of the essay I will discuss these several organizational levels by which this ultimate causal connection is implemented. Each one preserves and builds upon this basic capacity and extends its causal power into larger and larger domains.

## Beyond derived intentionality

My argument hinges on distinctions between different types of things that are sometimes termed "intentional", and in particular on the distinction between "derived intentionality", as discussed in detail by Dennett, and the intrinsic nano-intentionality upon which I focus in this paper. Biological systems can profitably be thought of as artifacts that have derived intentionality, but where the goal or purpose derives from the past history of evolution of the species. The oak tree doesn't *know* that it is about (say) photosynthesizing, resisting wind and storms, and obtaining adequate light and nutrients, so as to produce acorns and ultimately young oak trees. But these are nonetheless in some clear sense "goals" of the oak tree, necessary to explain many other detailed aspects of oak tree biology if we adopt (as we must) an intentional stance towards the oak tree. A crucial tool of evolutionary biology to adopt the intentional stance and utilize this functional way of thinking. But this is a statement about empirical science, and how we learn and understand: it is an epistemological and not an ontological argument.

However, in addition to this derived intentionality, most organisms have an additional type of "aboutness" that, once recognized, is different from, and goes beyond, the teleonomic "goals" of evolution (*sensu* Williams 1966). Organisms actually have *individual* goals of their own—such as individual survival, feeding, mating and healing themselves—that are *not* specific "goals" of evolution, but simply means to the end of the ultimate evolutionary goal of differential reproduction. Here, an over-ready adoption of the intentional stance can mislead us. For while we all know that (ontologically) the only causal power in the process of evolution by natural selection is differential survival, it's easy to slip nonetheless into thinking that Mother Nature "knows" about running speed, digestion efficiency and the like, because we (as intentional-stance-adopters) know about such things. But of course she doesn't—there is no one out there evaluating engineering efficacy except us scientists and (implicitly) the individual running, digesting organisms. A specific cell, following a chemical gradient in search of nutrients or mates, has its own individual-specific and temporally-bound local goal. Such intrinsic, local, individual goals become incorporated into multiple levels of cognitive complexity as we progress from unicellular to multicellular organisms, to micro-intentional animals with a simple nervous system lacking representation but nonetheless processing information and instantiating goals of the individual. They are further incorporated into complex fully-intentional organisms like ourselves which have nervous systems capable of representation and hypothesis formation.

## Evolving nano-intentionality and its derivative capacities

The thesis I am advancing depends on a hierarchical cascade of causal powers, of abilities that are characteristic of living things arranged in increasingly complex and differentiated systems. In this section I will combine an description of what these specific levels of causal abilities are with a historical discussion of roughly when and why they came about in evolution. This choice—to fuse phylogeny and function—will hopefully serve the dual expository purpose of making clear what the causal capabilities of these different levels of organization are, and of removing any whiff of mystery from how and why they exist on our planet. None of the biology discussed here is particularly controversial; although our fossil record of the early evolution of single-celled life is poor we are fortunate to have many exemplars of these ancient life forms among us today, within reach of any schoolchild with a microscope, eye-dropper and a local pond (Patterson and Hedley 1992). Many living single-celled forms (e.g. foraminiferans, radiolarians, shelled amoebae) produce tiny inorganic skeletons or "shells" that are important components of the fossil records, essentially identical in form to those found in today's oceans (Sen Gupta 1999). Thus we can be relatively confident in our assertions not only that some rather complex and impressive abilities typify single-celled organisms today, but also that they were typical one billion years ago before multicellularity appeared on our planet.

The thrust of my argument is, first, that the nano-intentional ability of cells in general to rearrange their structure in response to their circumstances represents a basic, primitive type of goal-directed "aboutness" that predated neurons, brains and minds. Second, this nano-intentionality is a *necessary* prerequisite for the mental "aboutness" that occurs when such a cell's circumstances involve information, as occurs when the cell in question is a neuron that is a part of a nervous system (Brown et al. 1991). The individual cell's rearrangement of molecules in response to an "information economy" like that created by the brain is the crucial bridging function between the physical and the mental. This ability is precisely what makes it possible for the "mental world" to have causal powers over the physical world: and vice versa. One might thus say that neuronal nano-intentionality is my substitute for Descartes' pineal gland. But as we will see, the connection between the causal power I call nano-intentionality and full blown intentionality is far more complex and involves several layers of additional organization. Although I suspect that this causal power is a logical necessity, in what follows I will be content to argue that it is just a brute biological fact: if you want mental intentionality, the way at least we vertebrates have it, you need nano-intentionality, the way eukaryotic cells have it. Whether a logical possibility of some other path to intentionality exists, for example in life forms on some other planet, will not concern me.

## "Aboutness" 101: nano-intentionality in an amoeba

*"Take what you have gathered from coincidence"*
*Bob Dylan, "It's All Over Now, Baby Blue"*

The most crucial step for my argument is already bridged by protists: single-celled eukaryotic organisms like an amoeba or paramecium (Grell 1973). An amoeba changes its structure, moving about "seeking" nutrients, engulfing food particles, following chemical gradients in a purposive fashion (Sandon 1963). Such behaviour provides the prototypical example of what I'm calling nano-intentionality. There is no mystery about how such physical changes occur: the process of cytoskeletal adjustment is based on a relatively well-understood process of random microtubule expansion and selective chemical stabilization (and can already be roughly understood in physical, mechanical terms) (Kirschner and Mitchison 1986; Kirschner and Gerhart 2005). Nor is there any mystery about where this "purpose" comes from: Darwin solved this problem for us. It is a simple statistical fact that those amoebae who rearranged themselves in certain ways survived longer and produced more copies of themselves than those who did things in other ways. So we can already cash out the origin of this simple form of nano-intentionality with no residual mystery. It is "derived" in precisely Dennett's sense. Furthermore, I trust that no one feels a need to attribute "mind" to the amoeba, because, being a single cell, it has no dedicated information-processing machinery even vaguely equivalent to a vertebrate nervous system (c.f. Dennett 1996). I don't think it makes much sense to wonder "what is it like to be an amoeba" since an amoeba has no separate "mental" system that exists to represent its body or its environment (Fig. 1).

Despite therefore lacking any *subjective* sense of purpose, the amoeba has a purposiveness that is undeniable, realized via its physical form: it is a complex arrangement of matter serving to do useful things like find food and avoid toxins. The amoeba can cope with novelty, and by changing its individual structure express new behaviour that is locally adaptive. It is this built-in "aboutness" that I am terming "nano-intentionality", and it is intrinsic to the cell. This purpose developed functionally and historically through the process of evolution and descent with modification, and its non-random usefulness is explained by natural selection. Although we don't yet understand exactly how this nano-intentionality is implemented, I don't think there are any residual *philosophical* questions to be explained in understanding this purposiveness: the remaining questions are for cell biologists. The purpose of an amoeba is instantiated causally in the physical structure of

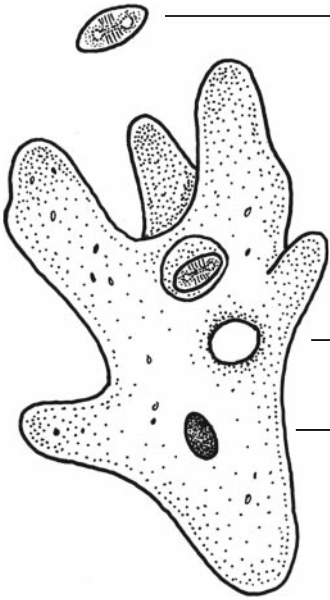


**Fig. 1** *Amoeba proteus* ingesting an algal cell—a prototypical example of a free-living single-celled eukaryote

cell membranes, organelles, cytoskeleton and the wealth of accumulated past experience stored both in these physical structures and in the informationally-encoded DNA.

These everyday characteristics of the amoeba or any similarly complex protozoan make it, not just qualitatively different from a transistor or a thermostat, but different in a way that provides a necessary building block for the very aspects of the vertebrate brain that we *do* find mysterious: its flexibility (specifically its ability to respond to novel circumstances), its adaptiveness, and ultimately its ability to generate subjective awareness and consciousness. The crucial pre-mental properties of a cell are that it can (1) respond to (somewhat) novel circumstances, eventualities for which it is not specifically-prepared by the evolutionary "memory" instantiated in its DNA, (2) discover, through an individual process of trial and error, some "adaptive" (in the physiological sense) response or solution, and (3) in various ways incorporate the results of this discovery into its own structure, thus "recording" or "remembering" (in a non-mental sense) this past, individual history.

Point one above reflects the fact that living things can sometimes adapt themselves to novel circumstances; despite the limitations on this adaptability it goes beyond the homeostasis of a thermostat or *E. coli* confronted with lactose. Point two describes the means by which this is achieved: via a selective algorithm formally equivalent to natural selection but played out at the individual microstructural level. A goal, random flailing and a satisfaction criterion can allow an individual to solve a problem for which its evolutionary history did NOT specifically prepare it. And point three is crucial because it shows how this individual history can be preserved for use in that individual's future tribulations. The information stored in the cell's 3-D structure is individual-specific. It dies with the cell, and cannot be passed on in its DNA. From an engineer's point of view, this is the magic step where the individual's arrangement of matter changes, autonomously, into something better able to cope with the future.

One might well ask where we draw the line between organisms that only possess derived intentionality (e.g. a virus) and those with the additional nano-intentionality. I have used an amoeba as an example because large, complex eukaryotic cells clearly have the capacity under discussion. I frankly do not know whether a bacterium such as *E. coli*, which lacks a cytoskeleton and organelles, and cannot crawl, engulf food, or deform its rigid cell wall, does. I am no bacteriologist, nor am I interested in drawing lines of this sort, so I will stick with examples of eukaryotic cells, which illustrate the fundamental issue unambiguously as free-living single-celled organisms. However, it is important to note that regulatory flexibility as the well-known switchable sugar metabolism made possible by the *lac* operon and similar systems (Jacob and Monod 1961) is *not* nano-intentional in my sense, because all of the capabilities and strategies are already there in the cell, directly coded in the DNA. There is little flexibility or capacity for individual novelty in this system, and its aboutness is entirely derived from the evolutionary process of random variation in millions of cells, combined with differential survival of certain variants. So, at least this aspect of bacterial behaviour does not pass muster as nano-intentional; but perhaps others do.

Before we ascend to the next level of complexity, it is worth reflecting on the fact that most of the evolutionary history of life on Earth was played out at the single-cell level. For about 2 billion years, amoeba-level nano-intentionality was all there was. Selection has thus had plenty of time to hone and perfect this aspect of living things before the advent of multi-cellularity. But there is an important limitation on single-celled organisms: a bacterium or an amoeba must be a jack of all trades, itself capable of everything involved in living: finding and catabolizing food, avoiding toxins, and reproducing. This constrains the overall form and function of single-celled organisms in some pretty obvious and significant ways, and is both the reason that amoebae look and function essentially the same now as they probably did one billion years ago, and why it took so long for life to get beyond this level. The way out of the "master-of-none" bind was multicellularity.

## Building on nano-intentionality: multicellular organisms and cellular specialization

When cells began to team up in multicellular complexes (a crucial evolutionary transition still re-enacted today by social amoebae like *Dictyostelium* (Kessin 2001)), the door was opened to specialization and the evolution of new cell and tissue types. For instance, the cells of a sponge (the simplest and least-organized form of fully multicellular life) are each genetically identical, and thus share each others' genetic interests 100%. This makes it stable, evolutionarily, for these cells to co-operate, and thus for each to specialize at some particular key task. Flagellar cells of the sponge are specialized to create water currents, skeletal cells secrete spicules to give the sponge its firm shape and create channels for the water to flow through, and amoeboid digestive cells pick out food particles, digest them, and share the nutritive results with everyone else. Each of these individual functions can be performed by single-celled organisms, but by specializing each cell type is freed from the "jack of all trades, master of none" constraint that plagues an amoeba or euglena (Fig. 2).

Crucially, this specialization of the nano-intentionality inherent in individual cells allows a composite intentionality of the sponge as a whole that is a bit (though not much) greater than the sum of its parts. A flagellar cell is "about" creating a water current, in a purposive way that a free-living euglena that beats its (physically identical) flagellum to

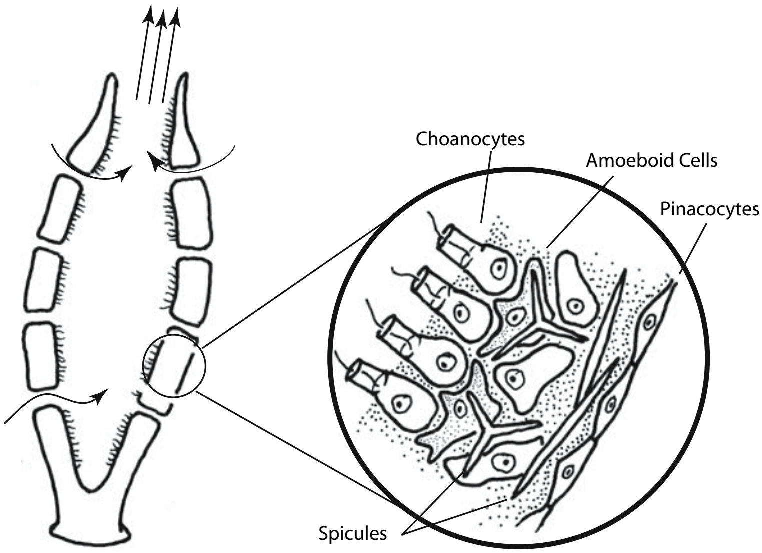


**Fig. 2** Schematic of a simple asconoid freshwater sponge (based on *Grantia*), illustrating specialization of cell types in this primitive animal. Inset: three cell types are illustrated: the *choanocytes* ("collar cells") set up currents and extract nutrient particles with their sticky collars; *amoeboid cells* move through the sponge absorbing and distributing nutrients; *pinacocytes* create the outer "body" layer, and glue it to the substrate. *Spicules* are non-living extra-cellular structural elements making up the "skeleton" of the sponge

locomote is not. Furthermore, sponges are "about" filter-feeding (that is, attaching to some surface, creating a rigid systems of tubes, creating flows through these tubes, and picking out the food particles that come through them) in a way that no single cell of the sponge, as an individual, can properly be said to be. Again, neither the origin nor ontology of this composite nano-intentionality is mysterious: it is reducible in materialistic, objective terms to the individual functions of its component cells, together with their 3-D configural relationship to one another within the organism.

Note, equally crucially, that the filter-feeding function of sponges is inherent in the sponge itself, to be understood in evolutionary terms, and is NOT dependent on an observer like a human postulating "the purpose of sponges is filter feeding". The observer-independent evidence is found in the various adaptive ways in which the sponge organism can rearrange itself to increase flows, and has specialized sub-behaviours that regulate and exploit flows. The sponge itself obviously doesn't *think* its "purpose in life" is filter-feeding, because the sponge itself doesn't think anything at all: none of its tissues are specialized for information processing in a way even vaguely recognizable as mental. Sponges are "about" filter-feeding because their ancestors were filter-feeders, and sponges better at filter-feeding (rather than photosynthesizing, detritus-feeding or various other functions) did better than those worse at filter-feeding. The aboutness of the sponge is *reflected* in its structure, but is *constituted* by its evolutionary history: the story of sponge-kind. This should not be too controversial, but it is crucial because the same logic will allow me to argue below that the aboutness of a brain is something inherent in the brain itself, as opposed to something imputed to it by an outside observer (one who has decided to adopt Dennett's "intentional stance"). This avoids an infinite regress, where the mindedness of the brain is dependent on some other mind inside or outside the brain observing the brain's activity (Ryle's "ghost in the machine" or Dennett's

"Cartesian theatre" of the homunculus). The brain does not need to "know" that it is about creating a mind, anymore than the sponge needs to "know" that it is about filter-feeding. Nonetheless, these functions are perfectly well-grounded ontological characteristics of the respective systems: both observer-independent and requiring neither an internal (subjective) observer, nor an external (scientific) observer.

So the crucial transition allowed by multi-cellularity is the possibility for specialization of the nano-intentionality inherent in the single cells, working together in a communal fashion. This specialization allows composite functions that would be impossible for a single cell alone. However, far from being left behind, amoeboid nano-intentionality was one of the crucial processes conserved in the transition to multicellularity, and such cellular behaviours remain crucial, highly-conserved processes in development of multicellular organisms like ourselves (Kirschner and Gerhart 2005). The next step in the logical (and evolutionary) path to mind is the advent of neurons: cells whose specialized nano-intentional purpose is a very unusual one: processing information.

## Micro-intentionality in simple nervous systems: evolving a proto-mind

A jellyfish, sea anemone or hydra has no brain, but it has a nervous system: an interconnected set of neurons arranged in a weblike fashion throughout its body (Ruppert and Barnes 1994). The causal powers of this "nerve net" allow responses to stimuli and control of body movements. Neurons are an important new type of cell; they are "about" amplification of information into locally adaptive patterns of action. "Information" is defined here as patterns of force and matter that are relevant to the purposes of the jellyfish: it is crucially jellyfish-contingent. In the same way that multicellularity allowed a specialization of skeleton and pump in the sponge, the advent of neurons constituted a wholly new form of specialization in these first true animals. This would be true even if there was one single neuron that (say) caused the jellyfish to contract when it fired in reaction to light, tissue damage, or some other stimulus. However, I know of no such organism. Nervous systems appear, from the beginning, to be implemented by groups of neurons (Fig. 3).

The nano-intentional purpose of a neuron in such a brain is to respond to and adaptively process information, directly analogous to the pumping or supporting functions of sponge cells. Its information processing (in terms of firing) is processed into physical, cellular changes in its shape, both at the micro-level of synaptic boutons and at the whole cell level of changing dendritic and axonal arbors (Brown et al. 1991). This is where the eukaryotic cell's general ability to adaptively change its form, using its cytoskeleton, becomes hijacked for specifically informational purposes. These direct causal relations of a cell in its nerve net, adapting to local stimulation, remain nano-intentional. Nonetheless, the *assemblage* of multiple neurons together leads to a wholly new level of composite intentionality (let's call it *micro-intentionality*) that is constituted by the relations of the nerve cells *to one another*. This is the crucial transition at which the causal powers that we assign to minds (rather than to bodies) become discernible. These relations provide the most elemental building block of mind per se, because now a specific event of neuronal firing can be properly said to be "about" something external in an immediate synchronic fashion, local to the organism (as opposed to the diachronically-derived evolutionary "aboutness" of the cell we have had up to this point).

The wave of firing in the nerve net, sparked by an external event, creates a higher order of causal efficacy, grounded in an active, moving body. The capacity for movement and agency at the whole-organism level, a response to the more active predatory life of

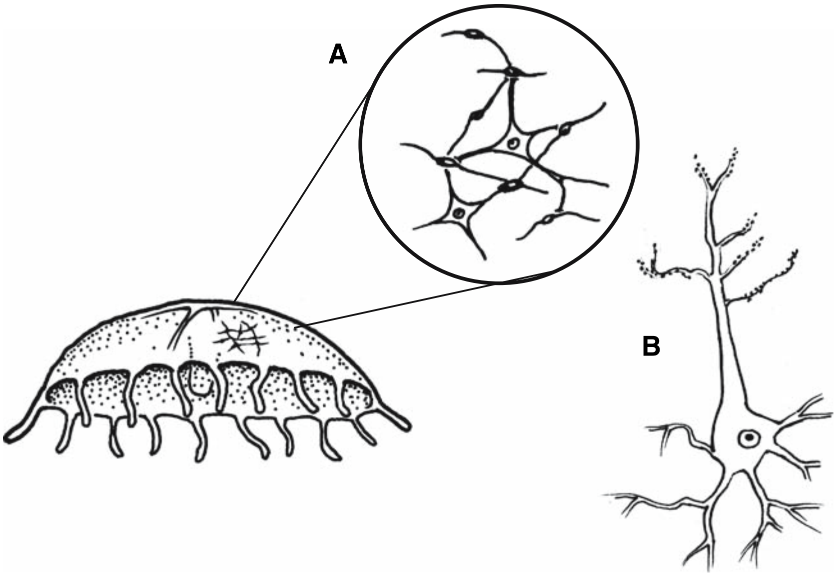


**Fig. 3** Medusa nervous system, schematically illustrating simple nerve nets. Jellyfish possess a nervous system made up of neurons very similar to those found in vertebrates. Inset A shows the nerve net, with two cell types (based on *Aurelia*). Inset B illustrates a mammalian pyramidal neuron from human cerebral cortex

coelenterates, means that the organism *as a whole* takes action. These actions then feed back down to the individual level of the cells (which must adapt to what are, from their viewpoint, community "decisions"). Thus, the possession of a nervous system adds to a body a new causal power, something beyond the mechanical nano-intentional firing of the individual cells, or the composite nano-intentionality of the sponge. The embedding of this cell-level event in a nervous system controlling a body makes it, at a very basic level, *proto-mental*. "Mental" because it instantiates the kind of causal relationship between events in the environment, including both perception and action, which is what minds are all about. "Proto" because many cognitive scientists might object to calling a system lacking representations "mental" and a jellyfish nervous system still doesn't have anything like a "representation" either of its own body, or of the outside world.

## Evolving representations: the dawn of mind per se

A dog's information processing capacities obviously dwarf those of a jellyfish in terms of versatility, flexibility and storage capacity, despite being made of neurons little different from those of the jellyfish. Advanced nervous systems, like those present in any vertebrate, are vastly more complicated than a coelenterate's nerve net, and thus the variety of composite, mental activity they are capable of supporting is far greater. The simple nerve net of a jellyfish exists in only the *actual* world. Although such a system might learn (making stimulus–response connections due to experience that can later be re-used) and can take novel action, what it can't do is generate a *model* of the world. A nerve net has no independent representations of the world that would allow this. In contrast, most of the neurons in more complex brains, like those of vertebrates, exist *only* to generate such models: "possible worlds" instantiated in patterns of neural firing. In a rat or dog or human brain, all of the neurons that are not directly sensory or motor neurons deal with information *as defined by their own nervous system* rather than with environmentally-derived information. Each of these neurons has its own history and purpose, developed via its own nano-intentional ontogenetic history. But it additionally functions in a second-order complex of information flows that is itself historically, ontogenetically contingent, at the level of the whole individual. An association neuron learns, not about the world per se, but about models of the world generated within the nervous system.

This is the place in my argument where micro-intentionality bridges into "true" mental intentionality, with an aboutness that is captured nicely by (say) the map of the retinal inputs that is laid out in primary visual cortex (V1). Once a primary representation is in place, so is the machinery for illusions and hallucinations, and more adaptively, imagery and imagination (e.g. Kosslyn et al. 1999). The nerve net of a jellyfish does not have such representations: and thus a jellyfish could not hallucinate or experience an illusion. The development of the representational system in V1 is contingent both upon nano-intentionality in the constituent nerve cells *and* upon a larger scale (mostly competitive) process of millions of cells interacting, constrained by visual input from the retina (we have reached the realm of ordinary textbook neuroscience so I won't belabor the point; see (Hubel 1988; Kandel et al. 2000)). By virtue of a combination of myriad micro-level cellular efforts, and whole-brain economic constraints characterizing the whole individual's ontogeny, V1 neurons are *intrinsically* "about" the retina in a quite meaningful and observer-independent way. This is because of retinotopy: the direct, physical connection, via a chain of cell bodies and axons, that isomorphically map specific regions of the retina onto successive stages of visual cortex. Visual neurons are directly, causally, connected to the photoreceptors they are "about". As many experiments in sensory-deprivation during development have shown, this developmental process of neural self-organization is highly contingent, and in no way guaranteed to occur, or pre-specified in the genes as a blueprint. In vertebrate development at least, this requires nano-intentional interactions among cells (including neurons, photoreceptors and glia (Purves and Lichtman 1980, Hubel 1988; Purves 1988; Brown et al. 1991). The intentionality in this system is thus *not* simply that which it inherits from the process of evolution, nor simply that of the individual histories of cells. It is a composite function of both, which builds upon them but adds an additional level of information and constraint. In such a nervous system, I think, it become sensible to talk about representation: the hallmark of the mental.

We are finally ready for the step of my argument: the final step leading to the dawning of a primitive form of awareness, of subjective "aboutness", in the history of life on earth. Although it remains unclear to me whether this next logical step actually mirrors an actual evolutionary step (perhaps the dawning of representations automatically and immediately leads to the next stage, perhaps not) I will keep things clear by stressing their potential independence.

## Consciousness: the evolution of serial awareness in a parallel brain

The last step to full-blown mentality has two logical steps, although in the nervous system itself (and perhaps in phylogeny) they are closely intertwined. First, we need representations (like the V1 map just discussed). But value can be added to such representations by other functionally separate component of the nervous system (the "associative" component of neuroscientists) which perform additional processing on raw representations, that tries out various alternatives and chooses among them. So this requires a further specialization of proto-mental intentionality into separate components: those that represent possibilities (and actualities), and those that process them. We have reached the level of organization that (Dennett 1996) dubs “Popperian”. Once a nervous system develops to a state of being able to *represent and decide among multiple possibilities*, the causal basis for awareness becomes necessary, and the final stage of my argument is reached.

The argument (as outlined previously in Fitch (2005)) starts with the curious fact that awareness is serial, even though information processing in the brain is massively parallel. Dennett expressed this odd fact nicely by saying that consciousness is like a “serial system implemented on a parallel device”, though as far as I can tell he doesn’t have a convincing explanation for why this must be so. I argue that this is a necessary design feature, because a functional body-control system (a brain) needs to respect a crucial constraint: there is but *one* body and one world, and often, only one chance to get things right. Nonetheless, a complex nervous system derives its power precisely from its ability to consider *multiple*, competing possibilities. Such a Popperian system entertains multiple possibilities but implement only one of them (its hypotheses can “die in its stead”). But, crucially, such a system requires some way of tagging and retrieving *the option actually chosen* if it is to learn from its successes and failures: it needs to know which neuronal assemblage is to blame when it makes an error, and which gets credit when things go well. Only thus can cells make their local, individual, nano-intentional adjustments to the whole-brain information economy: an individual neurons needs to be “told” about the composite actuality characterizing the whole organism of which they are a myopic part. Such a “tagging” of the actual versus the possible must necessarily be unitary (and thus serial in time), and a Popperian brain must have some way of implementing this tagging. The resulting causal power, I have argued, gives rise to the subjective sense of serial, unitary, purposeful *awareness* possessed by any organism with a brain capable of performing these functions (Fitch 2005).

Thus, finally, we reach the point at which nano-intenionality is directly linked to mental intentionality of the ordinary sort. For the stamping of neural events as realized vs. unrealized, and the assignment of credit and blame to neuronal assemblies is precisely where the aboutness of some global pattern of firing is integrated into the physical structure of the brain, and ultimately individual synapses, by nano-intentional neurons. This updating process is where mind (in its full, intrinsically intentional sense) makes literal, causal contact with matter, via the intervention of nano-intentional cellular reconfiguration.

To the extent that this argument is correct, awareness is the intrinsic subjective side of an objectively-verifiable capability of some types of nervous system to both entertain multiple hypothesis at a given time, and to later learn from their mistakes and successes. This leads to a rather strong and surprising thesis: that any nervous system objectively capable of considering and choosing among mental options, and of learning from its past decisions, will have at least a little bit of awareness as it does so: a little bit of (serial) consciousness. So yes, this thesis suggests the inclusion among the aware not just humans and dogs but also birds, frogs and fish, and probably some arthropods and molluscs. Note that I am obviously not talking about *self*-awareness—the suggestions that a grasshopper is aware that it is walking along my arm does not entail that it is aware of being aware (that it has any independent representation of itself as an entity). The number and complexity of parallel hypotheses considered will have an influence on the necessity for, and thus the degree of, awareness present. We too can be like zombies as we execute some highly over-learned task—or highly aware as we entertain a number of conflicting options.

**Summary of the argument**

In summary I have suggested a series of functional and evolutionary steps that allow us to move logically from the non-living to the living but non-mental, then to the mental but unconscious, and finally to the intrinsic intentionality of the mentally aware. Surprisingly, stage one is the most important step of my argument: the first 2 billion years of evolution on the earth led to eukaryotic cells which possess a crucially intrinsic aboutness I dubbed *nano-intentionality*. This was already present in single-celled organisms before the advent of multicellularity. In stage two, nano-intentional cells teamed up to form complexes—multicellular organisms—in which specialization of cell types could occur. In stage three, a new type of cell was born whose sole purpose in life is to process information: the neuron. As far as we know, neurons have always worked together in groups, as exemplified by the nerve net of a jellyfish or hydra, which I call *micro-intentional* by virtue of this information-processing specialization: the purpose of nervous tissue is to process information, in the same way that the purpose of a flagellar or heart cell is to pump fluid. I see this third stage as *proto-mental* because it lacks representations. The fourth stage simply expands upon the information processing of nervous systems by adding the capacity for *representations* of the body and the world—the hallmark of the truly mental. Finally, the necessity for choosing among various options represented, and "tagging" the one actually implemented, leads to the *serial awareness* that we subjectively experience as consciousness.

Each stage crucially retains, and builds upon, the intrinsic nano-intentionality of the individual cell. This allows me to arrive, in the end, at a true intentionality that is, in a specific implementation-dependent sense, intrinsic. Some of these steps are already well-established (e.g. from the non-representational to the representational) but there are two crucial novel steps. First, from a diachronically determined nano-intentionality present in all cells to a synchronic micro-intentionality present in the nerve cells in a multi-cellular organism like a jellyfish. The second step takes us from such a proto-mental organism, to one which can consider and choose among multiple hypotheses and learn from the experience, a process that entails a serial marking of the path actually chosen (which is experienced by the organisms as conscious subjective awareness). The other stages are straightforward biology and neuroscience. I think none of them should be empirically controversial, even if their philosophical implications remain so.

Although the framework that I've outlined is heavily steeped in the actual biology of life on earth, nothing in it stipulates that ours is the only way in which intentionality and conscious awareness could possibly be implemented. The whole argument stands (or falls) on the innocuous-seeming assumption that a single-celled organism like an amoeba has a "little bit of intrinsic intentionality" in a way that a thermostat simply doesn't, and that this intentionality is reflected in the cell's intrinsic causal ability to autonomously reconfigure its material form in response to its environment and its own individual history in a way that a thermostat or pulley or transistor can not. This is because the "intentionality" of a thermostat is not intrinsic to the thermostat, but is wholly borrowed (or better, imputed) by its human inventor. Although there are many steps following this core, crucial claim, the argument stands or falls mainly on this idea. Although we can make machines that "represent" and "choose" among alternatives and even "learn", by my argument they are not aware. The scare quotes are necessary, because truly *mental* learning and awareness are composite reflections of a more fundamentally and intrinsically nano-intentional capacity existing in the neural components of an organism's brain.

## Implications

The implication of my argument for artificial intelligence, in either the strong or weak forms, is pretty obvious. If a researcher aims to make a conscious machine, doing it with rigid switches (whether vacuum tubes or static silicon chips) is barking up the wrong tree. It would be necessary to go back to square one, technologically speaking, and create computing components and machines that can autonomously and adaptively reconfigure themselves physically, at a distributed molecular level, and then figure out a way to put a whole lot of them together in just the right way. I have no quarrel with the idea that the diachronic intentionality that was granted to us vertebrates by a long, slow process of evolution might be short-circuited by clever engineering design (by, so to speak, lending these machines some of our own long history of intentionality). But if the thesis I've put forward here is correct, the synchronic causal capability that is constitutive of mind will be difficult or impossible to implement on the static silicon chips or machines that are currently available, for brute physical reasons. Perhaps, to echo advice given to Dustin Hoffman in "The Graduate", "the future is in plastics". But my goal here is not to predict the technological future, but to simply describe current biological reality.

An obvious objection is that one could theoretically build a computer simulation of everything I've just discussed. Not with today's computers perhaps, but a desktop computer can already accurately model a gaggle of neurons in exquisite detail, down to the last synapse, and it won't be long before we can do the same for 100 billion neurons. Wouldn't *that* have intentionality and consciousness? This can be easily rejected as a criticism of the claim I am advancing, for a simple reason applicable to all computer simulations: that one can build an accurate model of something does not give that simulation the causal powers of the real thing modeled. So a synapse-accurate model of the human brain will not have all of the causal powers of a real brain, any more than a molecule-accurate model of the Gulf Stream can generate hurricanes or warm Scotland. As this difference between simulation and reality is both obvious and already noted by many (Searle 1992; Shapiro 2004) I won't rehash it. But for a committed functionalist, this argument seems to miss the point.

### Where functionalism went wrong

There exists a good reason to accept that neural simulations have different causal powers than weather simulations: if we accept that minds are *nothing but* information processing systems. Although this functionalist suggestion—that minds are algorithms independent of their implementation—is sometimes treated like an axiom of cognitive science (a view often, I think unfairly, associated with (Marr 1982)), I think it is a convenient fiction. Only if we unreservedly accept that implementation of an information-processing system makes no difference at all to its function, should we believe that an information-processing simulation possesses the same causal powers as the system simulated. As I hope to have shown, there is a very good reason to doubt this, in the case of vertebrate brains, because in addition to its information-processing role, an individual neuron has an additional set of much more ancient roles and causal powers, powers not present in gears, vacuum tubes, or transistors.

If we abandon the functionalist idealization, as I argue we must, we can begin to discern the circuitous evolutionary route to mind: from the nano-intentionality intrinsic to an amoeba to the concepts entertained by vertebrate minds, and the meanings that we express in the sentences of language. Although many promissory notes remain to be cashed out, I hope to have shown the value of pursuing such connections. At the very least, there is value in unpacking such monolithic notions as "intentionality" and "meaning" into more specific, biologically-grounded constructs. Even if some of the biological details are wrong, I think such details offer a way out of an otherwise knotty philosophical dilemma created by the computational, representational theory of mind. Cells (and thus brains) do more than "process information" and "represent"—they actually do work, and in particular they actively reconfigure their material molecular structure as part of their representing. At a fundamental microscopic level, neurons negate the mind/body dichotomy and all of the problems it causes.

## Original intentionality or original life: distinguishing life from mind

If nano-intentionality is a basic characteristic of most living species, could one not simply substitute the term "alive" for the causal power I have discussed in this paper? I think this equation is incorrect, for several important reasons. First, it seems likely that prokaryotes ("bacteria")—the most numerous living things on earth—*do not* possess nano-intentionality as I have defined it. Nature is fond of rejecting neat theories with inconvenient facts, and there is much left to learn about prokaryotic behaviour, but it therefore seems likely at present that most living things, numerically speaking, are not nano-intentional. Second, even if *all* living things possessed nano-intentionality, it would still be worthwhile to pick out the specific characteristic(s) that are relevant to the higher-level aboutness of biological information-processing. For me, the most crucial defining aspect of life is reproduction (autonomous self-replication)—something different from and far more basic than nano-intentional capacities. For example, vertebrate neurons are differentiated cells, unable to generate new cells, and thus have lost this most basic ability while retaining nano-intentionality. Finally, the deep formal similarities between the process of evolution by natural selection and the processes underlying thought have been repeatedly recognized (c.f., Maturana and Varela 1980; Bateson 1988; Dennett 1996; Godfrey-Smith 1996) and some scholars have even claimed complete equivalence between them (e.g. Maturana 1980). While accepting the value of such insights, I reject any direct equivalence of mind and nature. Unlike life, mind took a long time to appear on this planet, and we want to know why. There are many living things that do not have minds, and the neurons in our brains are among them. The origins of life, and more specifically nano-intentionality, are prior to the origins of mind, both logically and historically. Thus, the argument in this paper accepts a close link between the problem of life and the problem of mind, but clearly distinguishes between them. In doing so, it shifts a significant burden of explanation in the philosophy of mind from the origins and functions of brains, back to the origins and functions of eukaryotic cells.

**Acknowledgments** I thank Daniel Dennett, William D. W. Fitch, Phillip Pettit, Kim Sterelny, Gesche Westphal and an anonymous reviewer for comments and constructive criticisms of an earlier version of this manuscript, and Antonio Damasio for insightful conversations on this topic.

## References

Bateson G (1988) Mind and nature. Bantam Books, New York
Brown MC, Hopkins WG, Keynes RJ (1991) Essentials of neural development. Cambridge University Press, Cambridge, UK

Carroll SB, Grenier JK, Weatherbee SD (2005) From DNA to diversity: molecular genetics and the evolution of animal design. Blackwell Science, Malden, Massachusetts
Dennett DC (1987) The intentional stance. MIT Press, Cambridge, Massachusetts
Dennett DC (1996) Kinds of minds. Basic Books, New York
Fitch WT (2005) Computation and cognition: four distinctions and their implications. In: Cutler A (ed) Twenty-first century psycholinguistics: four cornerstones. Lawrence Erlbaum, Mahway, New Jersey, pp 381–400
Godfrey-Smith P (1996) Complexity and the function of mind in nature. Cambridge University Press, Cambridge
Goldstein ST (1999) Foraminifera: a biological overview. In: Sen Gupta BK (ed) Modern Foraminifera. Kluwer Academic, Dordrecht, pp 37–58
Grell KG (1973) Protozoology. Springer-Verlag, Berlin
Hubel DH (1988) Eye, brain, and vision. Freeman, San Francisco, CA
Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 3:318–356
Kandel ER, Schwartz JL, Jessell TM (2000) Principles of neural science. McGrawHill, New York
Kessin RH (2001) *Dictyostelium*—evolution, cell biology, and the development of multicellularity. Cambridge University Press, Cambridge, UK
Kirschner MW, Gerhart JC (2005) The plausibility of life: resolving Darwin's dilemma. Yale University Press, London
Kirschner MW, Mitchison T (1986) Beyond self-assembly: from microtubules to morphogenesis Cell 45:329–342
Kosslyn SM, Pascual-Leone A, Felician O, Camposano S, Keenan JP, Thompson WL, Ganis G, Sukel KE, Alpert NM (1999) The role of area 17 in visual imagery: convergent evidence from PET and rTMS. Science 284:167–170
Marr D (1982) Vision: a computational investigation into the human representation and processing of visual information. WH Freeman & Co., San Francisco
Maturana HR (1980) Biology of cognition. In: Maturana HR, Varela FJ (eds) Autopoesis and cognition. Reidel, Dordrecht, pp 1–58
Maturana HR, Varela FJ (1980) Autopoesis: the organization of the living. In: Maturana HR, Varela FJ (eds) Autopoesis and cognition. Reidel, Dordrecht, pp 59–140
Millikan RG (1987) Language, thought, and other biological categories: new foundations for realism. MIT Press, Cambridge, Massachusetts
Patterson DJ, Hedley S (1992) Free-living freshwater protozoa: a colour guide. Wolfe Publishing Ltd, Aylesbury
Purves D (1988) Body & brain: a trophic theory of neural connections. Harvard University Press, Cambridge, Mass
Purves D, Lichtman JW (1980) Elimination of synapses in the developing nervous system. Science 210:153–157
Ruppert EE, Barnes RD (1994) Invertebrate zoology. Saunders College Publishing, Fort Worth
Sandon H (1963) Essays on protozoology. Hutchinson Educational, London
Schopenhauer A (1819) The world as will and idea. JM Dent, London
Searle JR (1980) Minds, brains and programs. Behav Brain Sci 3:417–457
Searle JR (1992) The rediscovery of the mind. MIT Press, Cambridge, MA
Sen Gupta BK (ed) (1999) Modern Foraminifera. Kluwer Academic, Dordrecht
Shapiro LA (2004) The mind incarnate. MIT Press, Cambridge, MA
Tartar V (1961) The biology of *Stentor*. Pergamon, Oxford
Williams GC (1966) Adaptation and natural selection: a critique of some current evolutionary thought. Princeton University Press, Princeton, New Jersey