THEORY OF MIND

From Artificial Intelligence to Hybrid Intelligence

JOSÉ LUIS GONZÁLEZ QUIRÓS [1], DAVID DÍAZ PARDO DE VERA [2]

[1] Universidad Rey Juan Carlos, Spain
[2] Universidad Politécnica de Madrid, Spain

KEY WORDS	ABSTRACT
Philosophy of Mind *Contemporary thought* *Ray Kurzweil* *Artificial Intelligence* *Theory of Mind* *Hybrid Intelligence*	*Philosophy of mind has long ceased to be, if indeed it ever was, the exclusive domain of philosophers. In contemporary thought there is increasing interest in these matters from the point of view of technology. This paper gives a critique of the ideas of Ray Kurzweil and briefly reviews some of the recent trends in the treatment of these questions.*

GLOBAL KNOWLEDGE
ACADEMICS

One of the most common metaphors to speak of the mind is, without a doubt, that of the computer, or, more generally, the machine. This immediately shows that, at least in the sphere of folk psychology, we take for granted that thinking is an action and that, as with all actions, thought has an executant who uses an instrument or a tool —incidentally, we should remark that if we regard thought as an action it is because, in some way, thought take place through time, that is, because we are aware that time passes and we pass with it. Although we will not dwell on this observation, which we think is an essential one, we will try to extract some benefit from it later.

The image of the machine elicits a series of analogies that may have seemed useful and stimulating in order to understand that which is generally called the mind, especially when, since modernity, the Aristotelian notion of soul as the form of the body was abandoned. Although in its Cartesian origins the mind was primarily a conscious one entirely distinct from matter, very soon the view took hold that the thinking substance is just another product of the bodily machinery, and also relatively soon analogies began to be suggested that were supposedly enlightening as to its function. Whatever may have become of this idea, what we want to emphasize is that the philosophy of mind that has been predominant since the mid-20th century has assumed, very frequently, that the mind is a certain kind of device, a device that mediates strangely between an objective reality that is nevertheless in some way ungraspable, and a subjective reality that ultimately becomes objectivized in various forms or modes such as language, logic or science. Thus, somewhat surreptitiously, it was possible somehow to translate the *problem of consciousness*, which is rather intractable, into a sort of *problem of intelligence*, and to assume thereupon that if we could create a machine capable of thinking the way we seem to do it, there would be no fundamental objection to saying not only that the machine thinks, but that it has as much consciousness as one needs to be around in this world. In other words, in order to dodge the evident elusiveness of consciousness, the alternative taken was, as might be expected, to elude its study. It was then possible to speak of artificial intelligence and the mind-machine problem. The great advantage of this approach was that it presented scientists with a sufficiently large work programme to forget the reasons that might impel them to "waste time" with *metaphysical* enigmas. If, after all, it could be said that scholastic cosmology finally died at the hands of the telescope, it was reasonable to expect that *spiritualism*, to call it such, would end up dying at the hands of computer science. Philosophy of language was then to cleanse the mortal remains of the *ghost in the machine*, which is more or less what *eliminative materialism* has tried to do, as have in general all doctrines that have thought that an intensive verbal treatment[1] could remove from the horizon the *hard problem,*[2] as it has been called, along with other ones that tend to look easier than they really are.[3]

At any rate, there was an almost Freudian omission in this, let's say, programme; a very peculiar aspect of the matter was overlooked, a deficiency that has been pointed out recently by Roger C. Schank, [4] namely, that it is nigh impossible for us to imitate what we do, if we do not know how we do it. Indeed, because it is a big

---

[1] Something along these lines could be said also of those who insist on finding different models of approach, as if the problem could be reduced to a flaw in the logic of how it is formulated. We could regard as a suggestion of this sort the one made by Alva Noë (https://philosophy.berkeley.edu/noe) in a very interesting interview in Edge (which can be read at http://www.edge.org/3rd_culture/noe08/noe08_index.html), according to which the mind is not something that happens in the brain, because it is not *inside* of us, it is something we do.

[2] I believe David Chalmers (1995) and (1996) was the first to use this terminology.

[3] Putnam (1999:13) observes, for example, that "many philosophers want to dismiss traditional problems in the philosophy of perception as if too much time had already been wasted on them and as if we were simply beyond them now."

[4] "We can speak properly without knowing how we do it. We don't know how we comprehend. We just do. All this poses a problem for AI. How can we imitate what humans are doing when humans don't know what they are doing when they do it? This conundrum led to a major failure in AI, expert systems, that relied upon rules that were supposed to characterize expert knowledge. But, the major characteristic of experts is that they get faster when they know more, while more rules made systems slower". (https://www.edge.org/response-detail/11038.)

thing to forget that we do not know –in fact we have no idea– what it is we are doing when we think, or how we do it. Incidentally, this ignorance is not specific to thought, since we also have no any idea whatsoever –even though many people believe otherwise– what it is we do when we speak, or how we do it, or what we do when we move our heart, or when we breathe, although it is fair to admit that we seem to know a little more about the latter. In overlooking this, no doubt crucial, aspect of the matter, what the mind-machine programme did, in a sense, was to imitate the *product* of our action, not the *way to produce it*.

There is no good reason to brush aside what we might call a *primary* or, if you will, *naïve dualism*. It is quite useless to waste saliva trying to change the names of things in the hope of covering the problem or making it disappear, and there are no sufficiently powerful reasons that would impel us inevitably to accept some form or other of physicalism, be it physical, biological or digital. Nevertheless, we the authors have never had difficulty in recognizing that dualism is not a satisfactory solution either; it is just the best way to refer to a question, or family of questions, that are on the boundary of our view of the world, no less; a problem Schopenhauer named, brilliantly, the *riddle of the world*.

1. Artificial intelligence: faith in technology as a *cultural* factor, with a brief Ortegan *excursus*

When everything seemed to indicate that the failure of the prophecies by the first AI gurus had removed from the agenda the question of machine thought, one of today's great software creators has brought it back to life with his ideas, which not only renew the promises of what was called strong AI, but they complement them, concretise them and endow them with a rather peculiar halo of mysticism. We are referring to Ray Kurzweil[5], one of the best-known gurus of Silicon Valley (he was one of the creators of the so-called OCR programs and the software that can convert text to speech), who is convinced

that very soon it will be possible to attain a personal and bodily quasi-immortality (particularly bodily, through medication). Our man has agitated the scene with his publications, conferences, and even a documentary film about what he calls *spiritual machines* and the imminence of the perfect synthesis of minds and machines.

Kurzweil (1999:123) draws inspiration explicitly from the human genome project, launched in 1991, suggesting that just as it has become possible, at least in theory, to *map* the entire human genome, it will be possible to map the human brain, synapse by synapse. A wide range of very interesting possibilities will then lie open before us. For instance, we could save our personal memory in more robust files that would be immune to the well-known phenomenon of memory modification; we could *copy* our minds, aggregate them, erase them (partially, it is to be presumed); that is, we could do with our minds everything that we can do now in almost completely routine fashion with our computer files.

The connotations of Kurzweil's work suggest that we consider it from a broader perspective than that of theory of mind. If we leave the somewhat confined field of philosophy of mind, perhaps Kurzweil's ideas should be seen as a contribution to a current that has acquired some notoriety in the medical field known as *enhancement medicine* (or *enhancement technology*), which Juengst (1998:29) has defined as the set of "interventions designed to improve human form or functioning beyond what is necessary to sustain or restore good health." Technological enthusiasm is not a new phenomenon, but advances in the so-called NBIC domains (*nanotechnology, biotechnology, information technology, cognitive sciences*) have encouraged this sort of propositions, to the extent that enhancement medicine itself could well be superseded by what has already been called transhuman medicine or medicine for transhumanist enhancement (Wolbring, 2005). For the proponents of this new medicine, human physical and mental qualities are indefinitely perfectible. The yardstick of normality should no longer be the empirical mean but the "transhuman" state reached by virtue of applying

---

specific technologies[6]. From this point of view, Ray Kurzweil's ideas, according to which the moment is near when humanity will transcend its biological limitations and reach a symbiosis with machines, which in turn could be regarded as spiritual, would mean the elimination of all relevant distinction between nature, humanity and technology.

What is thus brought into play is a very nebulous and complex question, one that allows us to make an immediate connection between philosophy of technology and philosophy of mind. More generally, Putnam (1999:69) was absolutely right when he wrote that "a nice allocation of philosophical problems to different philosophical 'fields' makes no real sense. To suppose that philosophy divides into separate compartments labelled 'philosophy of mind,' 'philosophy of language,' 'epistemology,' 'value theory,' and 'metaphysics,' is a sure way to lose all sense of how the problems are connected, and that means to lose all understanding of the sources of our puzzlement."

In any case, we would like to observe that one of the first authors to reflect on the importance of technology to give meaning to human life was actually Ortega y Gasset, and he did so very brilliantly. Ortega was able to see in technology at least two very profound and interesting dimensions which he examined with great insightfulness, although with his characteristically wandering focus. On one hand he thought that technology could act as a gendarme of the spirit that subjects imagination and literature, its most frequent ally, to a regime of asceticism and continence. But, secondly, he also realized that technology evinced a capacity for invention that is part of the essence of man, and not merely a choice. In Ortega's analysis these two dimensions of technology meet at a relatively unstable point of equilibrium because fantasy plays an essential role in both: in the former, as a force that must be contained; in the latter, as the key to the meaning of technology.

Ortega (1996:114) wrote that, far from living on Earth, man actually lives on beliefs, on a philosophy, but in such a way that, in spite of it, he lives in a state of essential dissatisfaction, of maladjustment, perhaps because he is too aware of the limitations of his beliefs. And so, man starts out by constructing his own way of being in the world. Man's own way of inhabiting cannot be reduced to a kind of natural destiny; it is a discovery, a formula: this was the idea that Ortega presented in his lecture to the German architects in Darmstadt, a brilliant piece which nevertheless he bemoaned somewhat bitterly. He said there (1996:107), literally: "Man is, essentially, dissatisfied, and this -*the dissatisfaction*- is the highest quality man possesses; precisely because he tries to have things he has never had. This is why I often say that such dissatisfaction is like love without a lover or like a pain I feel in members I have never had."[7] This characterisation of man is debatable, but there can be little doubt that what Ortega asserts in some way explains both the behaviour of the *mass man* and that of the most demanding and eminent philosopher, typically unsatisfied with any theory.

Ortega had already used very similar ideas in *The revolt of the masses*, showing how the *mass man* experiences the absence of standards and coercion as an invitation to live his own way, to prevail (1962:180). "If the traditional sentiment whispered: 'To live is to feel oneself limited, and therefore to have to count with that which limits us,' the newest voice shouts: 'To live is to meet with no limitation whatever and, consequently, to abandon oneself calmly to one's self. Practically nothing is impossible, nothing is dangerous, and, in principle, nobody is superior to anybody.' This analysis constitutes the acknowledgement of an important historical change in which the technological civilisation has had a predominant role: in the past, life for the average person was fraught with difficulties, dangers and want of every kind, and this made it necessary to obey the law, to embrace deeply the observance of an extensive body of norms. But today's world presents itself to us -at least in appearance- as a safer, more abundant place, where we are not so compelled to respect social norms; where virtually no one challenges our right to live according to our own rules or our right to satisfy our desires, although it is surely

---

[6] José Luis Puerta and I have considered the medical aspect of these matters more specifically in González Quirós (2009c).

[7] Excerpts from Ortega (1962) translated for this paper from the original in Spanish.

the case that the passion to regulate, inseparable as it is from any form of power, has moved to other areas, endeavouring to conceal its presence in order to confound us more. This general impression that every individual has about the world he chances to live in eventually turns into an invitation to aspire, to lose his fear of the impossible.

Of course, Ortega was aware of the significance technology has in the fate of humankind, as well as the risks it can engender, and so he wrote (1996: 55), "perhaps the fundamental disease of our time is a crisis of desires, which is why all the fabulous potentiality of our technology seems as if it were of no use to us." Nevertheless, Ortega was able to see with great clarity that technology is not merely man's adaptation to the world, but, more properly, it is the creation of a new world, as, according to him, man is ill-adjusted to nature and desires a world of his own.

At the time of Ortega's death, it was impossible to foresee almost any of the advances which today we regard as decisive when considering the relationship of man with his environment and weighing the effects of technology on human society. However, he was able to see how a new and paradoxical horizon of confrontation between nature and technology was taking shape, when he wrote, "technology's victory aspires to create a new world for us, because the original world does not suit us, because we have fallen ill in it. The new world of technology is, therefore, like a gigantic orthopaedic device… and every technology has this marvellous and —as everything in man— dramatic tendency and quality: that of being a fabulous and great orthopaedic apparatus."

Returning to Kurzweil, we must put into today's context his promise that an unprecedented augmentation of our intellectual abilities awaits us, if we manage to coexist with those spiritual machines he speaks of, in a context in which, for example, we know there are orthopaedic legs that make it possible to run faster than with natural ones. Kurzweil believes that what technology can do with the body it will be possible to do also with the mind: a virtually endless improvement of its capacity and reliability, a reform of its glaring deficiencies, in

light of the abilities our self-knowledge allows us to dream of.

Kurzweil's assertion implies that in the field of minds and machines, of biological brains and computers, *Ortegan orthopaedics* will be vastly surpassed by a synthesis that will entail a veritable fusion, and, with it, the beginning of a new evolutionary phase for humankind. His position, which, aside from other opinions it may deserve, doesn't seem to see any difficulties beyond the technological, compels us to ask whether it is based on more than just a deliberately vague promise; on an extrapolation of dubious legitimacy.

We know very well that, at present, the human genome project has failed to meet even a small fraction of the hopes that some of its proponents aimed to arouse, and so, although it is not a good approach to tie an argument to a failure that may cease to be so,[8] we must question whether there is any solid foundation on which to support Kurzweil's promises. Kurzweil maintains that the power of ideas to transform reality is growing at increasing speed, and he proposes a theory he calls "Law of Accelerating Returns" to explain how, according to him, technology and the evolutionary progress it feeds behaves as an exponential function (2005:3). When we are able to understand how we understand, when we can turn our intelligence into an object and obtain its "source code", we will be able to review it and expand it in completely new ways. Human life will be transformed irreversibly (2005:7), we will be able avoid many fatalities, we will have immortality very close at hand, and, at the end of

---

[8] It is all too common to laugh at prophecies of the past when they have not come true, but this is probably not a very intelligent attitude. It is interesting to note that this happens both when things that it was thought would happen don't, and when very important things do happen that no one was able to predict. Regarding the former case, it may be interesting to review the list of prophesies collected by the late Arthur C. Clarke (1999: 536ff). Although it does not fit the latter case exactly, it is very interesting to take a look at the roster of changes (albeit only as far as the culture industries are concerned) that have happened in the last forty years, compiled by the editor Mike Shatzkin (https://www.idealog.com/blog/stay-ahead-of-the-shift-what-publishers-can-do-to-flourish-in-a-community-centric-web-world). That something has not happened so far is not always enough reason enough to believe it will never happen.

the 21st century (2205: 30), the nonbiological portion of our intelligence will be –he says literally– trillions of times superior to mere human intelligence without external aids.

Whoever wants to find precise reasons for such stupefying statements should not waste time looking for them in the works of Kurzweil. We at least have not been able to find them. Our author relies mostly on such claims as that technological progress has always been exponential, or that the small (?!) genomic difference between chimpanzees and human beings –which according to Kurzweil (2005: 5) is of the order of a few hundred bytes– has not prevented us from creating technological wonders. And, of course, he embraces, without a hint of doubt, the conviction that the computational notion of the mind is entirely correct.[9]

Although rivers of ink have been spilled about what is meant exactly by computational theory of the mind[10], the idea, in essence, is that there is a basic and illuminating analogy between how *hardware* and *software* are related and how the brain (or *wetware*, as it is sometimes called) and the mind are. At heart, the computational model of the mind, as Carver (2007: 101) has pointed out, embodies and reinforces the functionalist analysis, assuming that the black box of the functionalist model is a computer.

At the same time, we can see functionalism, in a sense, as a *consequence* of the behaviourist analysis of mental phenomena; a rather heroic endeavour to avoid the difficulties and paradoxes of the intuitive notion of conscious mind, which

some audacious philosophers, like Dennett, have not shied away from daring to undertake in order to eschew the threat of dualism. Thus, Dennett (1991: 430) has written: "if what you are is the program that runs on your brain's computer [...] you could in principle survive the death of your body as intact as a program can survive the destruction of the computer on which it was created and first run."[11] Kurzweil presumably goes beyond this assertion of Dennett's –although, as we see, he is not the first to do so– in assuming that we have something more than an illuminating theory available to us because, as his interpretation of Moore's well-known law predicts, we are on the verge of having technology powerful enough to allow us to move from hardware to the brain and software to minds, in the same way that we can move, biologically, from the mind to the brain and vice versa. To hold this view is to forget all the difficulties that philosophers have identified regarding such a conversion. It is surprising that it should be proposed to throw out of the window the body of problems that philosophers have analysed in this regard, without realizing how dangerous certain metaphors are when they are passed off as science. About "Functionalists who deny that knowing about the details of the brain will ever tell them *anything* useful" Francis Crick (1994:75) wrote that "[t]his point of view is so bizarre that most scientists are astonished to learn that it exists" (*Ibid*.). If we assume that, in the same way as, it seems, happens with neurons, we will make consciousness emerge when we find the organisational pattern that can make silicon think, we are saying something that can hardly be taken seriously. One of the fiercest criticisms of this idea comes from one of its very founders: Hilary Putnam. In one of his last contributions to this knotted question, Putnam (1999:86) maintains that, insofar as the

---

[9] This idea is not always shared by researchers in the field. We shall cite here the opinion stated recently by Noel Sharkey: "Roboticist Hans Moravec says that computer processing speed will eventually overtake that of the human brain and make them our superiors. The inventor Ray Kurzweil says humans will merge with machines and live forever by 2045. To me these are just fairy tales. I don't see any sign of it happening. These ideas are based on the assumption that intelligence is computational. It might be, and equally it might not be. My work is on immediate problems in AI, and there is no evidence that machines will ever overtake us or gain sentience." The full article can be accessed here: http://www.newscientist.com/article/mg20327231.100-why-ai-is-a-dangerous-dream.html?full=true.

[10] The discussion on the matter between Ray Tallis and Igor Aleksander (2008) can be viewed at https://link.springer.com/article/10.1057/palgrave.jit.2000128.

[11] We do not know if Dennet realises that this assertion could be taken, albeit only in some sense, as an updated version of the notion of soul as the form of the body, and that it may accommodate admirably well the idea, religious in this case, that the soul could very well survive the body, be eternal. We do not see much difficulty either, if we may continue with the diversion, in that a God capable of handling all the information (just a shade more informed than Laplace's demon) could restore our best bodily shape, in order not to leave us, so to speak, leading a ghostly existence.

computational property has not been given a determinate meaning, functionalism is science fiction erected on a misunderstanding. No one to date has managed to solve a poorly conceived problem, so the objection Putnam raises to reductionists can be equally valid for naïve approaches like Kurzweil's, who believes he can see clearly in the future that which he cannot understand now: "Saying 'Science may someday find a way to reduce consciousness (or reference, or whatever) to physics' is, *here and now*, saying that science may someday do we-know-not-what we-know-not-how" (1999:173).

Putnam further adds something particularly interesting: "Not only does rejecting reductionist pictures *not* entail abandoning serious scientific research but, in fact, it is those pictures that often lead researchers to misconceive the empirical problems" (1999:174). Putnam defends philosophy, which has all but disappeared from Kurzweil's writings, as in effect he embraces eliminativism, while not seeming to realise that not being a philosopher is not enough to avoid the problems, as, again in Putnam's words, "philosophical confusion reaches far beyond the studies of professional and amateur philosophers" (1999:175), and though "[m]any things deserve our wonder [...] the formulation of an intelligible question requires more than wonder" (1999:174).

2. On what Kurzweil seems to forget

Besides the abundant and recurrent criticism philosophers may have directed toward the so-called strong AI programme, its failure lied, essentially, in a glaring inability to create something like a living consciousness, and also in the very limited success achieved in imitating intelligent activities that early pundits thought were very simple. Although it may be somewhat unmerciful, I cannot resist citing again a few statements I included in another book [12] to characterise what I called *cyberphilosophy*. In its first steps, some of the more conspicuous proponents of AI proclaimed that their work could be regarded as, at least, the third greatest event in the history of humankind. Marvin

Minsky declared in LIFE magazine in November 1970: "In from three to eight years we will have a machine with the general intelligence of an average human being. I mean a machine that will be able to read Shakespeare, grease a car, play office politics, tell a joke, have a fight. At that point the machine will begin to educate itself with fantastic speed. In a few months it will be at genius level and a few months after that its powers will be incalculable." Despite the delay the plan had already in 1984, Roger Schank, who was usually more moderate and restrained than Minsky, asserted that one day there will be an omniscient machine, and that work was being done to this end. Statements such as these were probably what led David Gelernter to claim that the field of computer science is full of oddballs avid for novelty.

The programme certainly failed, but this did not lead to abandoning the underlying intentions, as Kurzweil's appearance on the scene shows. In May 2009 Kurzweil presented two documentaries about his ideas in Brookline, Massachusetts.[13] Karim Gherab was there, and he told me Kurzweil could only answer with vagueness when asked how he thought he might copy a mind (or a brain, if he felt this would be easier) on a computer[14].

In the book that was published containing a debate between Kurzweil and some of his critics, the only philosopher present was Searle (2002:71-72), who reiterates his well-known views on the matter, namely, the irreducibility of semantics to syntax and his "Chinese room" experiment (which, in my humble opinion, is a restatement of Leibniz' famous *mill* argument[15]), and he throws at Kurzweil three decisive objections. Searle finds fault firstly with Kurzweil's trying to give the public the impression that he understands what he actually doesn't understand, an ugly habit no doubt, and secondly with his taking as definitively established truths

---

[12] The corresponding citations are in González Quirós (1998), pp. 110 and 111.

[13] *Singularity Is Near* and *Transcendent Man*, presented at The Coolidge Corner Theatre on 11th May 2009 (https://www.kurzweilai.net/kurzweil-to-present-sneak-previews-of-excerpts-of-singularity-is-near-and-transcendent-man-films-at-coolidge-corner-theatre-may-11).
[14] He also resorted profusely to the peculiar and invalid argument that since techno-sceptics had been wrong many times, now he would be right.
[15] Monadology, 14.

theories that are not so; and lastly, Searle contends that we lack minimally clear knowledge of how the brain does what it does.

Michael Denton, a biochemist that leans toward the notion of "intelligent design", criticises Kurzweil for overlooking some essential characteristics of living beings that in no way seem to exist in the domain of machines. The first of these differences is the ability to self-replicate that living beings have at many different levels. This is a property that computers obviously lack, albeit some theoreticians, futurists as they are, have maintained that it will be possible to equip them with functions similar to this property life has, which is one that seems to establish a decisive difference.

Another property cited by Denton is the ability of living beings to grow and transform themselves, changing their shape and structure. All of this can certainly be called *information*, but we are very far from understanding and from truly knowing how to do what a simple hen does warming an egg: to unfold the information contained in the yolk so that a chick is developed. (This example is mine, not Denton's.)

Denton turns to Kant and his *Critique of Judgment* [16] to remind us of the peculiar entanglement of causes and effects that is typical of lifeforms, which does not seem reducible to an orthodox causal analysis from the point of view of science. The organic form is not susceptible to simple reduction. Denton (2002, 94) claims, for instance, that "no artifact has ever been built, even one consisting of only 100 components (the same number of components in a simple protein), which exhibits a reciprocal self-formative relationship between the parts. This unique property […] is the hallmark of organic design".

The readiness to jump over life, as if life were something simple, is markedly characteristic of the curious audacity of the functionalists and thinkers of Kurzweil's ilk. Life isn't easy to understand, if we mean understanding as we do in the sphere of mechanics. It would be foolish to deny that in the future things may be understood which we do not understand at all today, but it is no less arrogant to assume that we can disregard life if we want to comprehend something like

consciousness. Furthermore, life is a phenomenon that is strictly tied to time, it is temporary in nature and it has certain, let's say, *negentropic* qualities, to use the term introduced by Schrödinger, which should not be overlooked. Nothing is gained by assuming that all complexity can be understood from the same simple elements we already know, completely, it would seem. It is no dishonour to admit there are things that are, for now, beyond us. The heart of the mistake that we make when we disregard the qualities of life that seem to be irreducible to mechanics and computing is to confound the abstract with the concrete. Smullyan, who believes the identification of the abstract with the concrete is one of the most tragic philosophical errors of our time, has said that the mind is as concrete as an entity can be (Smullyan, 1984). The fact is that, up to now, and setting aside computational or literary fantasies, the only judicious assumption we can make is that minds can be found in living beings. Life is a very difficult phenomenon to understand. It is supposed to fall under the purview of biology, but, as Emilio Cervantes[17] puts it, life slips like water through the basket of science. It would be pertinent, at this point, to recall Leibnitz' caution [18] on the difference between the works of God and those of man, but let us move on. The mind, for its part, is temporal and intuitive, utterly singular and elusive; one almost feels the temptation to say it is the only thing we know concretely, something surely very different to what the most complex and creative of programmes could be like.

3. Science and the properties of minds

From an epistemic point of view, it may be said that the main advantage of primary dualism over any form of reductionism is that the empirical

---

[16] Let us recall that Kant maintained there would never be a "Newton […] of a blade of grass" (Kritik der Urteilskraft, § 75).

[17] This is the motto of his interesting blog: http://weblogs.madrimasd.org/biologia_pensamiento/
[18] "For a machine made by human artifice is not a machine in each of its parts. For example, the tooth of a brass wheel has parts or pieces which to us are no longer artificial things, and no longer have something recognizably machine-like about them, reflecting the use for which the wheel is intended. But the machines of nature, namely living organisms, are still machines even in their smallest parts, *ad infinitum*. It is this that constitutes the difference between nature and artifice, that is, between divine artifice and ours." (*Monadology*, 64, translation by Nicholas Rescher).

problem of the relationship between the brain and consciousness makes perfect sense for the dualist, whereas it is virtually nonsense for the reductionist positions. Those who believe there exists something different from the brain the neurophysiologists see, [19] a mind that is distinguishable from the physical body, the brain, in this case, know very well that there is something beyond their mere belief or the coherence and appeal of their ideas; they know very well that all of it is generated in a fiendishly complex system of associations and influences between conscious and temporal perception and the completely physical events that occur in the brain. Schrödinger (1992:93) made this point masterfully: "The world is a construct of our sensations, perceptions, memories. It is convenient to regard it as existing objectively on its own. But it certainly does not become manifest by its mere existence. Its becoming manifest is conditional on very special goings-on in very special parts of the world, namely on certain events that happen in a brain. That is an inordinately peculiar kind of implication, which prompts the question: What particular properties distinguish these brain processes and enable them to produce the manifestation?"

In addition to pointing out the metaphysical paradoxes of the case, Schrödinger outlines here an entire work programme which, in one way or another, is being carried out. This extraordinarily complex universe of associations is a matter for science, not philosophy, to explore, but it is not an impossible task, or one that does not make sense, however difficult it may seem today in our eyes.

Our current knowledge is that the human brain has an average of 86 billion neurons, and a "typical" neuron is connected through synapses to 1,000 to 10,000 other neurons.[20] The brain can establish or interrupt roughly a million connections per second, and it can keep usable information for decades, *labelling* it, using its meaning in manifold relations, changing its

location or modifying it when necessary, and while doing all of this it coordinates the work of hundreds of muscles and the processes necessary for the body to function, without making us aware of it. The brain can interpret thousands of signals correctly and make appropriate decisions in milliseconds. In addition, it allows us to think, speak, maintain relationships and learn. All this activity is being studied with technologies of increasing subtleness, and we are acquiring immense amounts of information that must be combined, evaluated, interpreted and put into a coherent theory. Today we know relatively well which parts of the brain are involved in perception, how the brain works, how it processes the signals it receives, how it forms memories and how it controls muscle movements. We know which regions are activated with speech, when we look at something or when we do simple calculations, and we are beginning to know what happens when we make decisions.

It would be unusual if the task on the scientists' hands were not to become complicated by categorial concerns. The fact is that advances in such questions are much slower and limited in significance than would be desirable. We have celebrated years, decades and, almost, centuries of the brain; no one doubts the importance of these matters, and yet there are no palpable signs of any change of paradigm, except in areas where prophecy prevails over science.

Some may be tempted to believe that this is a field in which, as a revived Kant might say, it is not possible to make progress along the sure path of science. This is not what I think, if I am allowed the immodesty of expressing my opinion. I prefer to believe that we may be at the gates of some truly spectacular advances, rather than to assume we will come to the end of science. This notwithstanding, I do not believe that the kind of progress we can expect will be capable of solving any metaphysical enigma, for reasons very similar to those that make me doubt we will be able to remove the veil of Maya or to travel beyond space and time.

¿What can we expect? As the true *empiricist* I would like to be, what I anticipate is that a more precise science than the present one, but not necessarily very different, will allow us to

---

[19] The brain of which Bertrand Russell (1995:186) said, "I have been taken to task for saying that what a physiologist sees when he examines another man's brain is in his own brain, and not in the other man's".

[20] From Eric Chudler's web page at University of Washington: "Brain Facts and Figures" (https://faculty.washington.edu/chudler/facts.html; accessed on 12th January 2021).

understand better how the brain works, and consequently open the possibility of contributing purposely to the efforts to enhance the intellectual abilities within our grasp. This will make it possible, for example, to create new instruments to develop our capacities for perception, and even our intelligence. In other words, I believe that the brain will have *exoinstruments* connected to it in a rather simple and effective manner, to enhance its performance. I expect some sort of intellectual orthopaedics, of hybrid intelligence, will be possible, and that this will open new paths. I presume, furthermore, that this new form of hybrid intelligence will not come solely from the side of hardware, but also from software, based on the foreseeable improvement of the system of signs we use to think and to calculate, and of the ways to automate their relationships through new networks that are external to us or, in some sense, also hybrid. All of this might seem like science fiction, but perhaps we are not so far away from it.

Science is beginning to break free from conceptual frameworks and images that were, in a way, constraining it, and to equip itself with methods that will allow small but solid and steady advances[21]. To explain better what I mean I will make use of an analogy, referring to biochemist Michael J. Behe's commentary on Darwinism, as I believe it is an appropriate metaphor for what I am trying to say. Behe's criticism of Darwin, or, more properly, of contemporary Darwinism, moves along various fronts and, naturally, it does not deny biological evolution across different aspects of the very broad phenomenon of life. What he calls into question severely is its explanatory value. For Behe, from the moment in which it was possible to open the *black box* of molecular biology, the kind of explanations at the organism level that are typical of Darwinian arguments lack all power. Since any sufficiently expressive quote demonstrating how Behe's makes his argument would be too long, I copy in a footnote[22] one of

---

[21] Terrence Sejnowski, a computational neurobiologist at *Salk Institute*, recognises this: "The way that neuroscientists perform experiments is biased by their theoretical views", https://stage.edge.org/response-detail/10881.
[22] To Darwin, vision was a black box, but after the cumulative hard work of many biochemists, we are now approaching answers to the questions of sight. The following five paragraphs give a biochemical sketch of the

eye's operation. [...] Don't be put off by the strange names of the components. They're just labels, no more esoteric than *carburetor* or *differential* are to someone reading a car manual for the first time. [...]
When light first strikes the retina a photon interacts with a molecule called 11-*cis*-retinal, which rearranges within picoseconds to *trans*-retinal. (A picosecond is about the time it takes light to travel the breadth of a single human hair.) The change in the shape of the retinal molecule forces a change in the shape of the protein, rhodopsin, to which the retinal is tightly bound. The protein's metamorphosis alters its behaviour. Now called metarhodopsin II, the protein sticks to another protein, called transducin. Before bumping into metarhodopsin II, transducin had tightly bound a small molecule called GDP. But when transducin interacts with metarhodopsin II, th GDP falls off, and a molecule called GTP binds to transducin. (GTP is closely related to, but critically different from, GDP.) [...]
The above explanation is just a sketchy overview of the biochemistry of vision. Ultimately, though, *this* is the level of explanation for which biological science must aim. In order to truly understand a function, one must understand in detail every relevant step in the process. The relevant steps in biological processes occur ultimately at the molecular level, so a satisfactory explanation of a biological phenomenon – such as sight, digestion or immunity – must include its molecular explanation.
Now that the black box of vision has been opened, it is no longer enough for an evolutionary explanation of that power to consider only the *anatomical* structures of whole eyes, as Darwin did in the nineteenth century (and as popularizers of evolution continue to do today). Each of the anatomical steps and structures that Darwin thought were so simple actually involves staggeringly complicated biochemical processes that cannot be papered over with rhetoric. Darwin's metaphorical hops from butte to butte are now revealed in many cases to be huge leaps between carefully tailored machines – distances that would require a helicopter to cross in one trip.
Thus biochemistry offers a Lilliputian challenge to Darwin. Anatomy is, quite simply, irrelevant to the question of whether evolution could take place on the molecular level. So is the fossil record. It no longer matters whether there are huge gaps in the fossil record or whether the record is as continuous as that of U.S. presidents. And if there are gaps, it does not matter whether they can be explained plausibly. The fossil record has nothing to tell us about whether the interactions of 11-*cis*-retinal with rhodopsin, transducin, and phosphodiesterase could have developed step-by-step. Neither do the patterns of biogeography matter, nor those of population biology, nor the traditional explanations of evolutionary theory for rudimentary organs or species abundance. This is not to say that random mutation is a myth, or that Darwinism fails to explain anything (it explains microevolution very nicely), or that large-scale phenomena like population genetics don't matter. They do. Until recently, however, evolutionary biologists could be unconcerned with the molecular details of life because so little was known about them. Now the black box of the cell has been opened, and the infinitesimal world that stands revealed must be explained.

his texts (1996: 18ff.) on the biochemistry of vision, which shows how far our understanding of this function has reached beyond the organ-level of analysis. Darwin, in any case, could not have reasoned any other way, as the discipline of biochemistry did not yet exist in the second half of the nineteenth century.

Post-Cartesian materialists have been materialists not because they know what the brain does, but in spite of not having the faintest idea of what it does and how. They have been *a priori* materialists, believers in the curious metaphysical idea, if I may put it this way, that one is more credible than two[23]; that it was reasonable that in the universe there should be a single substance, and that the supposition that there might be two is scandalous. In other words, for classical materialism the brain could have well been made of solid wood. This materialism emanated from something entirely independent of the peculiar biological nature of the brain. Here Behe's critique of Darwinism becomes meaningful. We do not yet have all the keys to understand how the brain works, but we have more every day, and it would be regrettable that exceptional research should be interrupted or squandered in the name of a metaphysics that would make us think we already know what mental states are, namely, the products of the theoretical brain, let's say, of a wooden brain.

I am perfectly aware that the analogy I have just drawn has, as all analogies do, serious flaws, and that materialism is slightly subtler than the assertion that the brain could well be a piece of wood. Nevertheless, the *black box* theories of the brain certainly have led to the absolutely unjustified rise in prominence of the computational notion of mind, which, as we have discussed, was developed on the foundation of the behaviourist and functionalist criticisms of cartesian dualism –which, incidentally, was utterly misunderstood–, acting, as it were, in a manner befitting the old Spanish proverb, *a moro muerto*, *gran lanzada*.[24]

---

[23] This is one of the objections that Sherington (1984) makes to materialism in the first half of the twentieth century.

[24] Literally, "*valiant lancing of a dead Moor*," meaning it is easy to claim a great victory over a foe that is already dead (as Falstaff would agree).

Finally, we cannot continue confusing the logical analysis of the products of the mind, or perhaps we may say the brain, with the true knowledge of what the structures of the brain are that make possible our conscious activity, and how these structures operate.

The confusion of the mind with a computational system has lasted for too long, so it is worthwhile to try to understand the reasons that have allowed such a woolly hypothesis to remain on its feet. I believe one of them is the fact, as forgotten as decisive, that, as Schrödinger liked to say, the mind can only be *experienced* in the singular, which allows us to speak in the fullest sense of the problem of *other minds*. Hence, this absolute singularness and indistinctiveness of minds can be accommodated easily in the computational model, for which the mind is a single abstract entity. I think there is a further reason that explains the confusion, if indeed it is a confusion, between mind and brain from this perspective. I am referring to the fact that the extraordinary technological advancements of the digital age have brought about a synthesis of reductionist technologies (or technologies inspired in knowledge that is reductionist methodologically) and the actual digital technologies, which are not reductionist or physicalist, but rather, they rely on the capacity to handle semantic properties, and that this synthesis has contributed to the confusion of the metaphysical meaning of the former and the latter. It is necessary, however, to distinguish both sources of technology. Doing so may help to dispel the illusion on which the computational model of the mind rests. Reductionism as an explanatory strategy seeks to determine the elements that make up a given reality and connect them to elucidate the phenomena that characterize it. On the other hand, the creation of any digital entity is a process in which signs or units are extracted from a meaning that is given previously; a meaning that is not deciphered but, *au contraire*, encoded digitally. That such encoding lends itself wonderfully to lightning-fast electronic treatment should not confuse us regarding this essential difference.

Entirely aside from any computational metaphor, the science of the brain is in a position to obtain ever more precise knowledge of how

neurological biology works, and we can expect this will lead to technology-based improvements of our intellectual performance, which without much hesitation we might call hybrid intelligence. I have no doubt these advances will, when the time comes, require we refine the theoretical paradigm that can accommodate all these new forms of knowledge, but I believe the most appropriate metaphysics will continue to be, in one way or another, *dualist*, and I would wager that no empirical discovery will ever knock down a well-established metaphysical category, such as the *mental* category is.

To conclude this analysis, I believe categories of this kind are sufficiently protected from any reductionist threat for various reasons of principle, such as those that link consciousness with perception of time, an aspect of reality that is not so fully at our disposal as space; or those that have to do with the peculiar freedom our understanding must enjoy, if we are to profess having the capacity to produce forms of knowledge that are not absurdly contradictory. Regarding this last consideration, one of its best and briefest articulations is Epicurus' caution – which in my view is as brilliant as the Euclidean proof that there are infinite prime numbers – that "He who asserts that everything happens by necessity can hardly find fault with one who denies that everything happens by necessity; by his own theory this very argument is voiced by necessity." [25]

I therefore completely agree with a recent text by Juan Arana (2009:302-303): "Since a light cough can interrupt my freedom and my very self-awareness, what problem is there in admitting that exercising it is linked to the coordinated or unaligned activity of some set or other of neurons, or to the emission and reception of this or that neurotransmitter? It will be providential when such mysteries are solved, if indeed their unveiling proves to be a benefit to humankind. However, and since what is sought on the horizon of the neurosciences is to establish a correlation between neuronal activity and the awakening of consciousness or the exercise of the will, the possible discovery of such correlations, even if it is complete, would in no way equate to a total naturalization of consciousness and freedom. The

only thing that would be discredited with such advances is a stark dualism of a kind that Descartes himself never defended. The unity of man implies that the dimensions body and soul we find in him are so intimately intertwined that we wouldn't know how to separate them fully, neither physically nor conceptually. Traditionally, it is assumed without any trouble that this affects our notions of soul, psyche or spirit, but the truth is it affects bodily or material notions in the same way. In other words, it makes as little sense to speak of mere neurons, molecules and atoms as it does to speak of mere consciousness, will or freedom. In both cases these are abstractions, aspects of reality that are more or less defined, which we have separated conceptually, artificially cutting their links with the whole of which they are a part, and assuming that we can use them as a first approximation to the truth we are searching for." [26]

4. Hybrid intelligence and the new collective mind

When attempts have been made to imitate how the brain works by building neural networks, the mistake has been to assume that a real brain uses the same kind of relatively basic physical and logical architecture as that of ordinary software. Research into the properties of synapses and how they work, however, suggests they are considerably more complex. The efforts of mathematicians, neurobiologists and many other specialists are beginning to show that the circuitry of the brain relies on compositional and physical properties which, possibly without the need to resort to quantum mechanics, will allow us to improve our knowledge of how the brain works, and at the same time inspire ways to build faster and more powerful computing architectures.

As regards possible advances in the knowledge of the *physics* of the brain that supports our mental activities, we will relate the story of the memristor, following the suggestions made in a paper by Justin Mullins (2009). In 1971, Leon Chua[27], a young electronic engineer in California,

---

[25] Epicurus, Aphorism 40 of the Gnomologium Vaticanum.

[26] Excerpt translated for this paper from the original in Spanish.
[27] Chua's very technical text can be found here: https://ethw.org/w/images/b/bd/Memristor_chua_article.pdf.

had the impression that the state of the theory concerning the nature of electronic circuits was mathematically deficient. In the same way as Mendeleev was able to suggest there were gaps to be filled in his table of chemical elements, Chua thought that in addition to the three known elements (or devices) in a circuit – the capacitor, the resistance and the inductor – there must be a fourth. The reasons for his belief were strictly mathematical in nature, namely, that between four entities (in circuit theory they are: the electric charge; its behaviour in time, that is, the current; the magnetic flux; and the voltage) that hold binary relationships with each other directly, there must be six basic ways in which they are related to each other. However, circuit theory posited only five, [28] which to Chua was a displeasing flaw. There was a missing relationship between the electric charge and the magnetic flux. This sixth relationship should do something that could not be achieved through combinations of the others. Specifically, Chua thought that it should behave like a resistor which changes its resistance when charge flows through it, and in addition *remembers* the amount of charge that has flowed through it by keeping its last resistance value when the current is turned off. Chua called his device a memristor (*memory resistor*). For almost forty years the memristor was a strictly theoretical creature, because no physical device existed capable of performing that precise function.[29] At the dawn of the new century, there began to emerge materials and designs that could do so. These were systems that worked at the nanotechnological scale, unobservable at the scale of the millimetre.[30] The discovery promised to be useful to build flash or solid-state memories, which deliver much faster writing and re-writing speeds. But, more germane to our purposes, the real novelty came when the truly surprising behaviour of a unicellular creature, the *Physarum polycephalum*, was described. It appears this slime mould could solve certain basic puzzles and, more

importantly, it seemed capable of anticipating events that occurred periodically, which led researchers to conclude that, despite it not having neurons, it must have some kind of memory. Max di Ventra, a San Diego physicist who knew of Chua's ideas, learned of the case and compared it to the behaviour of a memristor, and later, in collaboration with Yuriy Pershin (2008) and others, created a model of a memristor which, it seems, behaves as a synapse does.

We have mentioned this example because in our view it shows one of the many ways in which it will be possible gradually to advance our knowledge of the singular physics of the brain. We believe the key will be in that our technologies learn to do what living cells, what neurons, do, and more specifically how synapses work, for instance, without supposing, on the other hand, that our information processing machines can provide any insight. To the extent that this may be feasible – and the challenge is almost of extravagant difficulty – perhaps it would be possible to think of a convergence of minds and machines, which today is very distant from our immediate agenda. The nature of the superiority of life over formal design is such that it clearly refutes Kurzweil's (2005:478) supposition that "The patterns are more important than the materials that embody them." We do not mean to deny the need for a general theory of the functioning of the brain, for something which we may call a *model*, but we do maintain, firstly, that the groundwork of knowledge laid so far is still glaringly insufficient to propose such a theory, and secondly, that the models inspired in the formal analysis of knowledge, the purely logical and/or functional models, are based on an error of principle. One of the latest things in neurobiology has been to assume that the brain works in a Bayesian manner, through anticipation mechanisms, drawing particularly from Friston's research (Huang, 2008) and from popular science works like Hawkins' (2004).[31] But even supposing that this theory is fruitful, the most interesting knowledge will always be that of the properties

---

[28] Charge with current and magnetic flux with voltage give us two of those possible relationships. The other three, to make the total of five, are those that describe the three elements or devices considered traditionally in circuit theory, namely, resistor, capacitor and inductor.

[29] Indeed, the validity of the memristor even as a theoretical circuit element has been challenged. (See Abraham (2018).)

[30] See Strukov et al. (2008).

[31] A review of his book, as well as helpful information to follow Hawkins' ideas, can be found in
http://www.uoc.edu/uocpapers/3/dt/esp/climent.pdf

of neurons, and the systems they form, which make possible the Bayesian behaviour.

We do not think the as yet remote possibility that by these, let's say, empirical routes we may develop technologies that complement and/or enhance our mental functions should entail a conceptual revolution, although we have no difficulty in admitting that some of the countless revolutions suggested by historians have been based on much less. After all, some may come to recognise that, as Nicholas Humphrey[32] puts it, "the 'Cartesian theatre of consciousness', about which modern philosophers are generally so sceptical, is in fact a biological reality", that our brains are the way they are for the very purpose of staging this mental theatre. Even from a very neutral perspective, or a purely pragmatist one, we should be interested, not so much in solving a problem that is metaphysical in nature, but in expanding our knowledge of mind and brain, so that we can enhance our performance and develop supports to handle many difficulties and problems.

A thinker as scantly mythomaniac as Freeman Dyson[33] has spoken, for instance, about the possibility of implementing some form of radiotelepathy, a direct communication of feelings and thoughts between different brains. This is not too different from that what we do when we speak over the phone, for example. To make radiotelepathy possible, according to Dyson, we would need to know how to convert neural signals into radio signals, build microscopic radio receptors, and also be able to read the mental meaning of neural signals, which is something Dyson does not mention. This is knowledge we do not possess, but it is not inconceivable to attain it. Dyson also speculates with the possibility of feeling what a bird feels when flying, or what a deer feels when shot, a proposition we imagine may seem as fanciful to Thomas Nagel as it does to us.

There is a second sense in which we can speak of hybrid intelligence, if we consider the changes that will take place eventually, as a result of the quest for what we might call the "new collective mind", that is, integrating into our working intelligence the new information and knowledge resources afforded by the growth of the Internet, by its specialization and by each person's capacity to combine it all into a coherent image of reality.

It seems clear that two sources of innovation can be deemed to have great power to modify our intellectual habits, and therefore the functions and dimensions of our operative intelligence of things, namely, digitalisation and the functioning of the network we know as the Internet. The first novelty lies in the fact that the information available has gone from being recorded on physical mediums, from being written in material documents, to being recorded digitally, in documents that are intangible but very easy to handle. This makes possible spectacular improvements in accessibility, transparency, economy and participation. A second big innovation is that reading can cease to be a purely passive activity, in that readers are able to leave a mark on what they read. This will increase greatly the amount of information available, although it also entails certain obvious risks. Thirdly, it will be possible to write under very different conditions to those of the past, with barely any scarcity of documentary support and with the prospect of obtaining responses to what is written almost immediately. Consequently, conditions will be absolutely ideal for debates, conversations and correspondence between scholars, which are the means by which modern science began to rise.[34] Lastly, the spectacular rise in the number of people who can participate in any debate makes it reasonable to consider, as Tim O'Reilly suggests, that ideas are becoming in themselves a particularly relevant form of social software. We should not forget that the biggest business on the Internet, Google, is based precisely on exploiting very smartly what people do when they use the Internet. It is a tool that exploits *social software* as a primary source.

We should bear in mind that, sooner rather than later, we will have the ability to handle with some ease massive collections of data relevant to any particular purpose, and that those collections will be constantly renewed since, to put it this way, accompanying the observed facts we will always be able to find the best reasons that have

---

[32] https://www.edge.org/response-detail/10428.

[33] https://www.edge.org/response-detail/10313

[34] See González Quiros (2009 a) and González Quirós & Gherab Martín (2009 b).

been discerned to support their validity. This will mean that everything will be continually re-edited, in a manner of never-ending and instant story that some may see as a terrible threat, if they are slaves to the idea that the most solid truths ought to be written in some inaccessible place, in bronze letters. Those of us who believe truth is in propositions and that these are, above all, immaterial, as immaterial as our consciousness, have reason to celebrate, because we will have at our disposal something which we may almost be tempted to compare to an inexhaustible source of wisdom – at least, that is, for those judicious drinkers who may delight in it.

The decisive influence writing or the printing press have had on human progress is a commonplace of cultural history which everybody accepts, and it is also very clear that the digital era is creating possibilities that are considerably more powerful and effective than the ones ushered in by those historical technologies. Nevertheless, it is immediately evident to anyone faintly familiar with the use of the Internet that social, cultural and institutional burdens are very seriously hindering genuine technological possibilities. After all, Bacon was right, and, since information is power, mere technological power is not enough to attain real power. But that is, of course, another question.

References

Arana, J. (2009). ¿Puede la libertad ser suplantada por elementos sucedáneos? Anuario filosófico, XLII/2, pp. 273-303.

Behe, M. J. (1996). *Darwin's Black Box.* New York: Touchstone.

Carver, M. (2007). Minds and Computers. An Introduction to the Philosophy of Artificial Intelligence. Trowbridge, Wilts: Edinburgh University Press.

Chalmers, D. (1995). Facing up to the problem of consciousness. Journal of Consciousness Studies, 2(3), pp. 200-219. Accessible from http://www.imprint.co.uk/jcs.html.

Chalmers, D. J. (1996). *The Conscious Mind, In Search of a Fundamental Theory*. New York: Oxford University Press.
Clarke, A. C. (1999). *Greetings, Carbon-Based Bipeds!* New York: Ian T. Macauley.
Crick, F. (1994). *The Astonishing Hypothesis: The Scientific Search for the Soul.* New York: Scribner.
Davidson, D. (2006). Mental Events. In E. Lepore & K. Ludwig (eds.), *The Essential Davidson* (pp. 105-121). Oxford: Clarendon Press.
Dennett, D. (1991). *Consciousness Explained*. Boston: Little, Brown & Co.
Denton, M. (2002). Organism and Machine: The Flawed Analogy. In J. W. Richards (ed.), *Are We Spiritual Machines? Ray Kurzweil vs, the Critics of Strong A. I.* (pp. 78-97). Seatle: Discovery.
González Quirós, J. L. (1998). *El porvenir de la razón en la era digital*. Madrid: Síntesis.
— (2009 a). El trabajo intelectual en el entorno digital: nuevas formas de escritura y erudición. *Arbor*, CLXXXV, 737, pp. 541-550.
González Quirós, J. L. & Gherab Martín, K. (2009b). Arguments for an Open Model of e-Science. In B. Cope & P. Angus (eds.), *The Future of the Academic Journal* (pp. 63-83). London: Chandos Publishing.
González Quirós, J. L. & Puerta, J. L. (2009c). *Tecnología, demanda social y medicina del deseo'. Medicina Clínica* [MedClin(Barc).2009. doi:10.1016/j.medcli.2009.07.002]
Hawkins, J. & Blakeslee, S. (2004). *On intelligence*. Nueva York: Times Books.
Huang, G. T. (2008). Essence of thought. *New Scientist*, 02624079, 5/31/2008, 198(2658).
Juengst, E. (1998). What does enhancement mean? In E. Parens (ed.), *Enhancing Human Traits: Ethical and Social Implications.* Washington, D.C.: Georgetown University Press.
Kurzweil, R. (1999). *The Age of Spiritual Machines*. New York: Penguin.
—(2005). *The Singularity is Near*. New York: Penguin.
Lepore, E. & Ludwig, K. (eds.) (2006). *The Essential Davidson*. Oxford: Clarendon Press.
Mullins, J. (08 July 2009). Memristor minds: The future of artificial intelligence. *The New Scientist,* 2715. Can be found at http://www.newscientist.com/article/mg20327151.600-memristor-minds-the-future-of-artificial-intelligence.html?full=true
Ortega y Gasset, J. (1962). La rebelión de las masas. Obras completas. IV. *Revista de Occidente.* Madrid.
— (1996). Meditación de la técnica y otros ensayos sobre ciencia y filosofía. *Revista de Occidente*.
Pershin, Y. V., La Fontaine, S. & Di Ventra, M. (2008). Memristive model of amoeba's learning. At http://arxiv.org/abs/0810.4179
Putnam, H. (1999). *The threefold cord: Mind, body, and world.* New York: Columbia University Press.
Richards, J. W. (ed.). (2002). *Are We Spiritual Machines? Ray Kurzweil vs, the Critics of Strong A. I.* Seatle: Discovery.
Russell, B. (1995). *My Philosophical Development*. London: Routledge.
Schrödinger, E. (1992)*. What is Life?, with Mind and Matter & Autobiographical Sketches.* Cambridge: Cambridge University Press.
Sherrington, C. S. (1940). *Man on his Nature.* Cambridge: Cambridge University Press.

Searle, J. (2002). I Married a Computer. In J. W. Richards (ed.), *Are We Spiritual Machines? Ray Kurzweil vs, the Critics of Strong A. I.* (pp. 56-78). Seatle: Discovery.

Smullyan, R. (1984). *5000 B.C. and Other Philosophical Fantasies*. St. Martin's Press.

Tallis, R. & Aleksander, I. (2008). Computer models of the mind are invalid, (text by Tallis followed by a discussion between both authors). *Journal of Information Technology*, 23, 55–62. doi:10.1057/palgrave.jit.2000128

Wolbring, G. (2005). The Triangle of Enhancement Medicine, Disabled People, and the Concept of Health: A New Challenge for HTA, Health Research, and Health Policy. Health Technology Assessment Unit, Alberta Heritage Foundation for Medical Research, HTA Initiative 23, December 2005.