5 Neuroscience
Dualism in Disguise

Riccardo Manzotti, Paolo Moderato

In this chapter, we have two goals. First, we want to highlight a surprising fact that is often denounced but seldom believed—namely that most of current neuroscientists, contrary to often-heralded physicalist credo, embrace dualism. Second, we want to introduce an original explanation of such a fact—an explanation that casts a disturbing light on many notions of current usage in the field of neuroscience. We will claim that the implicit assumptions adopted by most neuroscientists invariably lead to some sort of dualistic framework.

The observation that neurosciences are based on a dualistic conceptual framework is by no means new (Bennett 2003; Hurley and Noë 2003; Rockwell 2005; Uttal 2001, 2004). On the contrary, it keeps resurfacing with the same regularity with which most neuroscientists keep forgetting it.

One of the first to emphasize the intrinsic dualism of neuroscience has been John Dewey who claimed that "the older dualism of soul and body has been replaced by that of the brain and the rest of the body" (Dewey 1916: 336). Similar opinions have been expressed again and again (Gibson 1979; Holt 1914; Varela, Thompson and Rosch 1991; Whitehead 1925). In a strikingly similar note, Bennett and Hacker warn against a mutant form of Cartesianism lurking in neuroscience:

> It was a characteristic feature of Cartesian dualism to ascribe psychological predicates to the mind, and only derivatively to the human being. Sherrington and his pupils Eccles and Penfield cleaved to a form of dualism in their reflections on the relationship between their neurological discoveries and human perceptual and cognitive capacities. Their successors rejected the dualism [. . .] but the predicates which dualists ascribe to the immaterial mind, the third generation of brain neuroscientists applied unreflectively to the brain instead (Bennett 2003: 72).

In this regard, the philosopher Ted Honderich prefers to speak of *cranialism* (Honderich 2006a, 2006b). Recently, a very similar notion has been

further developed by Teed Rockwell who, in a provocative book, made the claim that

> Modern physicalists have kept the brain-body distinction even though they have thrown away the mind-body distinction, and are thus left with a philosophy of mind that is still in many ways fundamentally Cartesian: Descartes said the soul was in the brain, and identity theorists say the soul is the brain. Descartes' basic concept of mind is not really changed; it is simply demoted to being a concept referring to a particular kind of physical thing (Rockwell 2005: xi).

Hereafter, for coherence, we will make use of Rockwell's choice of words and thus we will refer to Cartesian Materialism to refer to a kind of disguised dualism. In this chapter, we will consider the various causes that determined this curious state of affairs and that lead neuroscience to embrace implicitly what it is, more often than not, publicly rejected. One of the main objectives of this chapter is to show that the standard premises of neuroscience invariably lead to a dualistic framework.

As a proof of the dualism pervading neuroscience, consider these two examples. The first is offered by Christof Koch when summarizing his life of research of the nature of phenomenal experience:

> Subjectivity is too radically different from anything physical for it to be an emergent phenomenon . . . I see no way for the divide between unconscious and conscious creatures to be bridged by more neurons. Experience, the interior perspective of a functioning brain, is something fundamentally different from the material thing causing it and that it can never be fully reduced to physical properties of the brain . . . I believe that consciousness is a fundamental, an elementary, property of living matter. It can't be derived from anything else; it's a simple substance (Koch 2012: 118–119).

So much for the reductionist stance professed in the title! The other example comes from the late Benjamin Libet who remarked "as a neuroscientist investigating these issues for more than thirty years, I can say that these subjective phenomena are unpredictable by knowledge of neuronal function" (Libet 2004: 5).

The structure of the chapter is as follows: First, we will outline the theoretical landscape in which neuroscience is trying to grasp the nature of the mind and that of consciousness. In particular, we will focus on the role of internalism. Then, in the next section, we will show the empirical and conceptual difficulties facing neuroscience when tackling with the issue of consciousness. In the third section, we will outline the main argument—namely that the empirical and theoretical obstacles, plus the internalist assumption, issue ontological promissory notes. In short, since it is impossible to locate

consciousness inside neural activity, scholars introduce fictitious entities as foundations of phenomenal experience. This is a preposterous form of dualism in disguise—publicly despised and privately practiced—that hampers any further understanding of the nature of the mind.

1. NEUROSCIENCE IMPLICIT ASSUMPTIONS

Neurosciences are not metaphysically innocent, as they would like to be. Although many scientists claim that their work is free from any undemonstrated ontological premises, this is not (and cannot be) the case. Any empirical data needs to be interpreted from the perspective of some premise. In this case, what are the assumptions on which most of neuroscientific research is based? Let us consider a quick overview of current neuroscientific discussion about the mind.

A first element is the supposed centrality of neuroscience in the study of the mind. It is a fact that a majority of neuroscientists and philosophers alike believe that only the study of neural activity will result in the explanation of the mind. Consider the title of recent book from a well-respected philosopher of the mind like Jessie Prinz: *The Conscious Brain* (Prinz 2012). The title itself reveals the central tenet of Prinz's work—namely that the mind is a property of the brain alone. The widespread consensus in science is that neuroscience is going to be the field that will finally provide a scientific theory of consciousness. The prevailing belief is that "the mind arises from the wetware of the brain" (Modha et al. 2011: 62). Neuroscience is proposing itself as the forthcoming *mindscience* (Manzotti and Moderato 2010).

Christof Koch fleshed out the gist of most neuroscientific approaches: "The goal [of neuroscience] is to discover the minimal set of neuronal events and mechanisms jointly *sufficient* for a specific conscious percept" (Koch 2004: 16, italics in the original). This is a rather precise claim: not only the brain as a whole is expected to produce the mind, but parts of the mind are taken to be the result of the activity going on in parts of the brain (Crick and Koch 1990). Yet, so far, this hypothesis has never been demonstrated empirically. Just to be clear, what is at stake is not whether neural activity has a role in enabling and tuning conscious experience but whether there is a given neural activity that is either sufficient for or identical with a given phenomenal experience.

A second element is that, in reality, neuroscience is mostly a physiological field of enquiry. Its methods are suited to study the activity of the CNS and its cells. In this regard, neuroscience has had incredible success during the last century. As of Golgi and Cajal's time, neurophysiologists unfolded the cellular foundations of our nervous system (Changeux 2001; Gazzaniga, Mangun and Ivry 1998; Kandel, Schwartz and Jessel 1991; Marijuàn 2001). The study of the nervous system has been carried out at all level of analysis and details, from the biochemistry of neurotransmitters up to the

computational models of huge cortical networks (Aizawa 2007). Currently, full-fledged models of large portions of the cortex are available, and many research groups are struggling to reproduce a complete working model of the cortex (Ananthanarayanan, Esser, Simon and Modha 2009; Modha et al. 2011; Sporns, Tononi and Kötter 2005; Sporns 2011; The Human Brain Project 2011). However, as amazing as these findings are, this is not exactly the same as studying the mind—an activity for which *neurosciences have not specific skills nor methods*.

A third element consists in a systematic confusion between physicalism and internalism—namely the hypothesis that if the mind is a physical phenomenon, it takes place as a result of neural activity. It is a view that the influential philosopher Jaegwon Kim stated crystal clear: "If you are a physicalist of any stripe, as most of us are, you would likely believe in the local supervenience of qualia—that is, qualia are supervenient on the internal physical/biological states of the subject" (Kim 1995: 160). It is a surprising confusion, since physicalism is the thesis that whatever the mind is it has to correspond to a physical phenomenon. However, this thesis does not entail in any way that the mind has to be *internal to the CNS*. Of course, if internalism would turn out to be true (and it is an empirical question), internalism would be a particular case of physicalism. For instance, behaviorism was definitely a physicalist view, although it was not an internalist one (Watson 1913; Hull 1943). We cannot rule out the possibility that there might be future forms of physicalism rejecting internalism.

It may turn out that internalism is false while physicalism is true. For instance, there are a handful of authors taking into consideration that physical constituents may be physically outside the CNS (Chemero 2009; Manzotti 2006, 2011; Rockwell 2005). It is surprising that so many scholars (Kim is one of the most notable examples) confuse physicalism with internalism.

A fourth element is a kind of historical inertia that neuroscience has accumulated during its development. For many neuroscientists, the final surrender of the mind to neuroscience is neither a philosophical thesis nor an explicit empirical hypothesis. It is just what they take for granted. It is so obvious, that it appears superfluous to state it. Neuroscience is based on a history of successes that originates in medical sciences.

Since the first autopsies at the end of the Middle Ages, medicine has had an enormous success by locating various phenomena inside the patient's body: from infection to blood pressure, from metabolism to movement, and so forth. As a paradigmatic example, consider the discovery of muscular strength inside the myofibrils by Luigi Galvani in the eighteenth century. He was able to make important progress because he was able to locate a function (movement and strength) inside an organ (muscles and myofibrils). The same explanatory template has been applied again and again. Many scholars are quite confident that in order to explain consciousness rather than idle philosophical speculation "a more practical approach is to use the tools of neuroscience that are available now to shed light on the neural

structures and activity patterns that underlie consciousness" (Tononi and Koch 2008: 239). As a result, in neuroscience, there are great expectations that the solutions to the problem of consciousness must respect the traditionally successful strategy—namely to single out a proper internal organ (likely the brain).

This successful strategy is still dominant. Like the stomach is the organ for digestion, like the lungs are the organ for respiration, like the heart is the organ for blood circulation, so the brain *has to be* the organ for the mind. At least, this is the prevailing expectation in science nowadays. Yet, it is just an expectation. It may turn out that the mind requires something totally unexpected—either new empirical data or a new conceptual twist. As we will see in the next section, against all expectations, neuroscience is still a far cry from naturalizing the mind.

To recap, the widespread beliefs in neuroscience may be articulated and synthesized in two independent and autonomous premises:

1) The mind is physical [DR, *dualism rejection*].
2) The part of the physical world which is sufficient to the mind is the brain (or some suitable proper part of the CNS) [NC, Neural Chauvinism].

The first premise is tantamount to a denial of dualism. The second premise is akin to what Alva Noë and Evan Thompson dubbed the "thesis of the minimally sufficient neural substrate" that is that "for every conscious state, there is a minimal neural substrate that is nomically sufficient (as a matter of natural law) for its occurrence" (Noë and Thompson 2002: 4). We won't discuss the first premise here. However, we will argue that neurosciences embrace both (and they don't have to) and that, by doing so, they are surprisingly compelled to adopt of form of camouflaged dualism.

As anticipated, these two premises together have the surprising consequence to develop a kind of dualistic picture of the mind in which there is, on one side, the world and, on the other side, the mind. This conceptual outcome is akin to the aforementioned Rockwell's Cartesian Materialism. However, it is not the end of the story. In fact, the same premise also compels neuroscientists to use a terminology that keeps separate the mental domain from the physical domains of neurons.

2. NEUROSCIENCE AND THE REPETITIVE FAILURE TO ADDRESS CONSCIOUSNESS

The previous two premises would not lead necessarily to dualism if it weren't because of neuroscience's failure to locate any phenomenon that may confidently be deemed as being identical with consciousness. In fact, the original formulation inside neurophysiology has been concocted in terms of identity theory (Feigl 1958; Smart 1959). However, identity theory was a failure.

According to Roger Sperry, "from the objective experimental standpoint, it is difficult to see any place in the material brain process for the likes of conscious experience" (Sperry 1969: 532). As result, the repetitive failure to substantiate the identity theory pushed many to withdraw from weaker forms of explanation such as emergence or correlation. Yet, identity theory was the original and natural choice for neurophysiology. It is a fact that any other scientific field of enquiry is based on identity theory. Consider genetic inheritance. Scientists have been looking for the obvious choice: some phenomenon that was the carrier of genetic information. In that case, the identity option was so obvious that it wasn't even discussed. But in the case of neuroscience it is far from obvious, very far.

In this section, we want to stress once again that, so far, neuroscience has been utterly unable to find any convincing physical phenomenon able to play the role of phenomenal experience. This incapability, which is likely the symptom of some wrong premise, has been eventually glorified as a feature of the mind-body problem. From a scientific perspective, there are only three viable options: the first is to admit failure, the second is some form of eliminativism, the third is denial. Apparently, failure is not an option, and eliminativism is no longer fashionable (an exception is, of course, represented by authors like McGinn or Dennett). As we will see, denial is the preferred choice by the scientific community. Denial takes the form of dualism in disguise.

As to consciousness, is the neuroscience situation really so desperate? We are afraid it is. For many years, neuroscience adopted a "don't ask don't tell" strategy with respect to consciousness:

> The already existing fields that study the mind or the brain have ignored consciousness. Psychology, behavioral science, cognitive science and cognitive neuroscience have avoided consciousness or have been reluctant to put subjective experience into the focus of their research programs (Revonsuo 2010: xxi).

Then, as for the '90s—thanks both to the enthusiasm triggered by new techniques for brain visualization and to the interests expressed by famous scientists—neurosciences discovered consciousness (Crick 1994; Edelman 1989; Jennings 2000; Miller 2005; Penrose 1989). The difference between mental processes and their alleged neural underpinnings results in heated discussion as to the causal role of neural activity (Lingnau, Benno and Caramazza 2009). Yet, this enthusiasm was not so successful as many hoped. In 1976, the neurophysiologist E. Roy John maintained that "we do not understand the nature of [. . .] the physical and chemical interactions which produce mental experience (John 1976: 2). In 1989, the psychologist Sutherland was still complaining that "[c]onsciousness is a fascinating but elusive phenomenon; it is impossible to specify what it is, what it does, or why it evolved. Nothing worth reading has been written on it" (Sutherland 1989).

Even Koch in his last book admitted that currently there is no clear model as to how neural activity becomes conscious experience (Koch 2012).

How can this situation endorse the enthusiasm that other authors seems to profess? Consider Prinz's statement that "these twenty-five years of inquiry have borne much fruit. Stepping back from this great mass of research, one can find various strands of evidence that point toward a satisfying and surprisingly complete theory of how consciousness arises in the human brain" (Prinz 2012: 3). Isn't this statement inconsistent with the perduring lack of results as to the physical nature of phenomenal experience?

The answer hides in a common misunderstanding. In fact, to a certain extent, Prinz is right. We know a lot more about the neural processes enabling conscious experience. Thanks to neuroscience, we know a lot more about the neural processes *involved* with arousal, sleep, memory, perception, free will, motor control, and imagery. Unfortunately, we don't know anything about the physical processes that *gives rise* to consciousness or, more poignantly, about the physical processes that *are* consciousness itself. This is a sort of confusion that hampers the discussion about consciousness in neuroscience.

One thing is to show that a neural process plays a role in tuning, enabling, and modifying a moment of consciousness. In this regard, any empirical evidence of this kind is scientifically of high interest, and it may possibly lead to a future breakthrough as to the nature of consciousness. But, *per se*, it does not tell anything about the nature of conscious experience itself. Consider the heating system in Jane's flat. Jane ignores whether it is the result of burning oil, burning gas, or electricity. However, in the basement, Jane discovered a control device that allows her to enable and tune heating. Jane discovers also that there are reliable correlations between the state of the control device switches and the resulting heating. Is the control device actually doing any heating? Of course not. Right now, most of neuroscientific data about consciousness are akin to the relation between Jane's control device and heating. There is a correlation, but scientists do not have a clue as to why such neural activity should result in conscious experience.

As a proof of such a lack of real progress between current neuroscientific data and theories and consciousness, consider this question:

> *Is there any neural activity that, at the best of our knowledge, may not happen without any consciousness?*

For instance, consider the activity in the fusiform gyrus that we know is strongly correlated with conscious perception of faces (Andrews, Schluppeck, Homfray, Matthews and Blakemore 2002; Kanwisher 2001; O'Craven and Kanwisher 2000). Let's skip the fact that the activity in the fusiform gyrus is not always correlated with conscious experience (Steeves et al. 2006). Even if fusiform activity and conscious perception were perfectly correlated (and they aren't), is there any reason why such a neural activity couldn't occur

without any consciousness? We are not aware of any. What would happen if we were able to replicate the same neural activity in a piece of neural tissue in a lab? On this possibility, Ned Block replied that he never heard anyone stating "that if a fusiform face area were kept alive in a bottle, the activation of it would determine face-experience—or any experience at all" (Block 2007: 482).

The same argument holds for all known neural activity. There are no compelling reasons why a neural activity may not occur without any correlated conscious experience. This highlights the big difference with other phenomena, which have been more thoroughly explained. Consider heat. Could we increase the average speed of molecules without increasing heat? No, that would be impossible because heat is the average speed of molecules. Alternatively, could we change the number of hydrogen ions in a liquid without changing its acidity? No, because we know what the relation between hydrogen ions and acidity is. Nevertheless, notwithstanding the impressive amount of collected data, there is no known neural process whose occurrence may not take place in the absence of conscious experience. Maybe there is. The point here is that we don't know of any.

If the mind is not identical with neural activity, there should be some kind of explanation of when, why, and how neural activity brings mental content into existence. So far, neuroscience is not even trying to do this. In the past, it mostly tried to debunk such puzzles by claiming they were ill-posed problems. This scapegoat is no longer acceptable. In order to define the mechanisms that generate a specific conscious experience, we need "to understand the conditions that determine what kind of consciousness a system has" (Tononi 2004: 1). Even if it were possible to identify a neural correlate of consciousness, why should a specific physical process lead to the occurrence of a specific phenomenal experience?

On the basis of which law? Currently, there are no psychophysical laws bridging the gap between physical and mental processes. What we ought to expect from neuroscience is some law of the form

$$\textit{Mental content} = F\ (\textit{neural activity})$$

Where F represents a law expressing a correspondence between neural activity and mental content. If anything like F were available, it would be straightforward to define a sufficient NCC (neural correlates of consciousness. For instance, the sufficient NCC of my conscious percept of red, C_{red}, would be the neural activity N_{red} occurring in my brain, such that $C_{red} = F\ (N_{red})$. Up to now, F is nowhere to be seen. F would endorse an internalist view of consciousness—namely that consciousness is something concocted out of the ongoing neural activity in the brain.

The main reason for the continuous failure of neuroscience to address consciousness is that the chasm between conscious experiences runs as deep as science itself. Modern science originated as a result of a crude

oversimplification—namely the separation between phenomenal qualities and quantitative properties (Galilei [1623] 1960). While the former has become hostage of philosophy and psychology, the latter have given rise to the impressive development of science. To recap, modern science expunged the conscious mind from its description of nature, and thus it obtained a simplified but very efficacious picture of physical reality. Yet, sooner or later, the chickens have to come to roost. How is it possible to deal with the mind once the mental aspects of nature have been methodically and, we would add, a priori eschewed from reality? It is like declaring that there are no forces acting at a distance and then trying to explain gravity. Something is not going to work.

Consider Prinz's definition of physicalism (2012: 11) as the "the conjecture that the fundamental laws and elementary parts that we find in things that lack mentality are the only fundamental laws and elementary parts in the universe". It is a surprising definition based on the mental domain admitting as physical entities only those "things that lack mentality". It is questionable for at least two reasons. First, this definition is parasitical on the mental. Second, it rejects a priori that the mental may be part of the physical.

To make a long story short, neuroscience cannot tackle consciousness because the essential properties of the mind have been programmatically and selectively set aside from the physical world. The list of such properties is not exactly the same although there is a consensus on the main ones: quality, unity, duration, intentionality, and first-person perspectives.

At present, neuroscience does not have a clue as to what is the relation between neural activity and consciousness. We learned a great deal on neural mechanisms that influence our conscious experience. Yet, nobody can predict for sure that a certain chemical activity is going to produce a phenomenal experience—not to speak of predicting what kind of phenomenal content. Of course, neuroscientists collected an impressive amount of evidence as to which neural activity is correlated to which phenomenal experience. It is enough to check the available literature either on anesthesia or on perception to see the extent to which we know the details of neural activity correlated with phenomenal experience (Mashour and LaRock 2008; Morimoto, Nogami, Harada, Tsubokawa and Masui 2011; Watkins-Pitchford and Brull 1997). But nothing in the literature explains why a certain neural phenomenon should produce a certain phenomenal experience. The evidence so far collected is just brute data.

Neuroscience faces an impossible mission—namely showing how a physical world which had been a priori defined devoid of those properties that are essential for the mind (unity, intentionality, quality, duration, and causation) may contain/produce those properties. This mission is taken to be somehow possible because inside the brain it is expected that something out of the ordinary may indeed happen. Yet, this would be mostly unexpected and indeed contrary to the starting premises. In this regard, J. J. C. Smart observed that

> There does seem to be, so far as science is concerned, nothing in the world but increasingly complex arrangements of physical constituents. All except for one place: in consciousness . . . So sensations, states of consciousness, do seem to be the one sort of thing left outside the physicalist picture, and for various reasons I just cannot believe that this can be so (Smart 1959: 142).

We do share Smart's concerns. It would be quite surprising, to say the least, to discover that the brain is the only place in the universe where awareness arises.

3. NEUROSCIENCE LEADS TO DUALISM: EPISTEMIC ENTITIES AND ONTOLOGICAL PROMISSORY NOTES

The rejection of dualism and internalism—together with the inability to single out a physical phenomenon identical with consciousness—determines three alternative outcomes: either eliminativism, identity theory, or dualism. The argument is as follows:

1) Neuroscience, for the reasons previously seen, rejects dualism and holds true internalism.
2) As a result, consciousness must be physically located in the brain.
3) So far, in the brain there is no empirical evidence of any phenomenon with the properties of phenomenal experience—namely intentionality, quality, unity, and first person perspective.
4) From a conceptual perspective, since the physical world is devoid of the aforementioned properties and since the brain is part of the physical world, in the brain there cannot be anything with such properties.
5) To conceal 2) and 3) and 4), neuroscience adopts an explanatory strategy that consists in presenting the mind as if it were in the brain. Mental properties are here but not like physical stuff. They are there, and they are not there.
6) That strategy consists in adopting a confuse terminology that makes use of ontological promissory notes such as code-talk, information-talk, computation-talk, and model-talk.
7) That strategy ends up concocting a form of dualism in disguise. They have the logical form of dualism, but they don't want to pay the ontological price for it.

As previously stressed, neither is internalism implied by physicalism nor is it the result of any empirical evidence. It is a hypothesis that got very wide acceptance both into the neuroscientific community and into the philosophical community. Inevitably, this hypothesis entails that something special is going on inside the CNS—something that contradicts the idea that the

physical world is devoid of mental properties. In fact, if the world is devoid of them, the brain, which is a part of the world, should be devoid too. Something conflicts.

Frankly speaking, the impression is that internalism's adoption and its inability to find any phenomenon akin to consciousness call back into service the ghost in the machine—albeit in disguise. Yet, this outcome runs afoul of the heralded rejection of dualism. This rather embarrassing situation resulted in entering into debt with ontology by ontological promissory notes that, so it is promised, will be paid back in the future.

In very loose terms, the sleight of hand is the following. The CNS is presented as the place where the mind ought to take place. Unfortunately, the mind does not look like anything material we may find *inside* the CNS. This is often quite embarrassing. Usually the blame is put on consciousness rather than on any potentially wrong premises. Anyway, at this point of the discussion, some other entity—which is not exactly a physical thing but that has some allure of scientific respectability like information, coding, computation, maps, representations, and symbols—is usually introduced in our description of what's going on inside the CNS to fill the gap. Since the hop from neural activity to consciousness appears disheartening, an intermediate stage is introduced. The intermediate entity plays thus the role of an ontological promissory note. It is something that is not really there, at least in material terms, but that nevertheless plays some role. After a suitable amount of discussion, the entity gain sufficient scientific respectability and epistemic prestige to be accepted as the special ingredient that will justify the appearance of the white rabbit of consciousness out of the brain-hat.

Consider information. Is information something more than the physical basis that implements it? This is unlikely, at least for a physicalist. Take my pocket calculator. If it rests on my desk, does it contain information? If it is still in the sealed box from the manufacturer, does it? Isn't it the same from a physical perspective?

Isn't it is just a piece of electronic junk that has a clever causal structure carved in its circuits? This clever causal structure is such that a human being may use the pocket calculator to do math. However, is there anything like information in addition to the causal and physical structure of the pocket calculator? We would rule out such an option because there is no physical mean to measure whether there is anything like information inside it. The pocket calculator may be described adopting an informational stance once a human being uses it to do math. Yet, a human being is able to do so in virtue of having a mind. Moreover, this brings us back to the usual problem of the mind.

Take a mark on a blackboard. Does it contain/bear/connect with a bit of information? How could we know it? We claim that there is no way to do it without knowing whether that mark is used by a human being to represent information. Since we assume that one may use that mark, we attribute to that mark a bit of information. However, information is not there like the

mass of the chalk that was used to mark the blackboard. We can check whether that mark, which is a physical object, has mass, charge, length, and so forth. We cannot check whether that mark has information. The same mark may be associated with 1 bit of information or with terabytes of information. It depends on whom and how its physical structure is exploited to connect external causal states of other systems.

In sum, the ontology of information is murky at best. Information is not equivalent with mass, charge, length, and the rest of physical properties. Without any pretense of originality (Floridi 2004; Landauer 1992; Manzotti 2012; Searle 1980), the intuition here is that information is just an epistemic entity. By epistemic entity, we refer to things like a center of mass that does not exist but that are useful concepts. We are not claiming that information is not an important scientific notion. We claim that a theory of information is something different from a theory of electricity. Thus, we stress the ontological lightness of such a notion when it comes to be the basis for further phenomena (consciousness included). Once again, we suggest that information (or computation, modelling, representing, and so forth) is like the notion of the center of mass. In fact, where the center of mass is supposed to be, there may be an empty space. The center of mass is a useful notion, which is the result of our mathematical description of gravity that benefits from assuming that the mass of a body, instead of being spread in a large volume, is concentrated into an ideal point—namely the center of mass. Useful as it is, the center of mass is nothing but an epistemic fiction introduced to simplify computations of gravity forces. The center of mass is particularly convincing because the body dynamics is such that, in many cases, their behavior is as if their mass were concentrated in their center of mass. However, it is only as if. The mass remains spread everywhere the body extends.

When an epistemic entity is used as if it were able to carry on ontological work, it becomes an ontological promissory note. There is nothing wrong in using terms such as centers of mass or information as long as they are taken to be ontologically empty. For instance, suppose that one develops a theory about dark matter such as that dark matter is explained in terms of centers of mass. The theory may be so difficult to verify that the focus on its complexity may distract from the simple fact that it doomed to fail since it is based on ontological promissory notes—that is, on nothing.

Here, we suggest that the hallmark of an epistemic entity is that i) it is something that cannot be ascertained in isolation and that ii) it does not add anything to the causal description of the world. Consider again the center of mass. There is no way to check whether a point in space is a center of mass just by inspecting the proposed location. Furthermore, one could dispense from using the very notion of the center of mass by referring to the actual mass distribution.

Similarly, any attempt to build a theory of consciousness based on vague epistemic entities is condemned to be unsuccessful. Consciousness is a real

phenomenon. It needs to be grounded on ontologically real entities not just on epistemic entities. Information does not appear to be the right building block.

A nice example of this questionable epistemic strategy is offered by the neuroscientist Ronald Melzack and his neuromatrix theory that "proposes that pain is a multidimensional experience produced by characteristic 'neurosignature' patterns of nerve impulses generated by a widely distributed neural network . . . in the brain" (Melzack 2005: 1378). Once again, we are faced with a vague notion that is neither mental nor completely physical. On top of the mere physical neural activity, we have patterns characterized by neurosignatures. Are these neurosignatures something more than the physical world? Because if they are just physical, they cannot have any mental features (like pain). If they are more than physical, then Melzack is advocating a form of dualism where the pattern/impulse dichotomy substitutes the traditional mental/physical dichotomy. Yet the logical structure remains the same.

A similar dichotomy appears in the work of most neuroscientists (see Table 1). Consider Haggard who is compelled to resort to a curious combination of terms such as "conscious awareness" as if there could by anything like unconscious awareness. Or consider the widespread use of terms like "interpretation", "coding", "mental states", and "mental content" (Haggard 2002; Haynes and Rees 2006; Kay, Naselaris, Prenger and Gallant 2008; Nishimoto et al. 2011), not to speak of other popular tools of trade such as "map" and "computations" (Li 2002; Roe, Pallas, Hahm and Sur 1990; Shagrir 2012; Wandell and Winawer 2011). We have no pretense to exhaust the literature here. Libraries have been written on it.

Consider another favorite example of neuroscientific jargon—modelling. Neural activity models reality. Many neuroscientists find scientifically acceptable the idea that the brain builds a model of the world. For instance, Thomas Metzinger suggested that the mind is the result of a multilayered

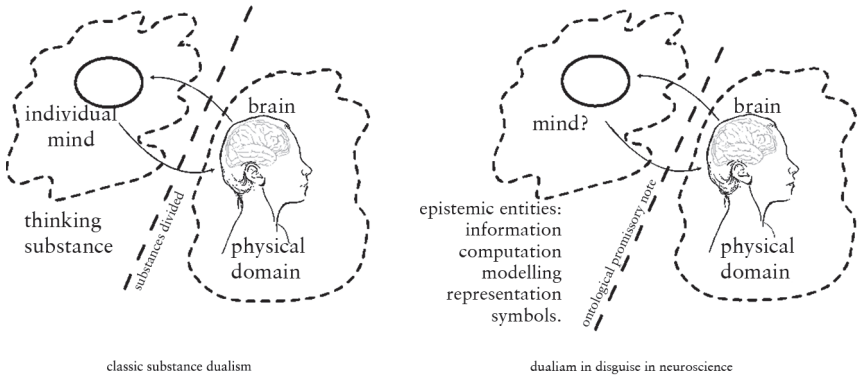

Figure 5.1

structure of models: "The subjective experience of being someone emerges if a conscious information-processing system operates under a transparent self-model . . . you constantly confuse yourself with the content of the self-model currently activated by your brain" (Metzinger 2003: 1). It is a revealing sentence that shows how an intermediate level (ontologically empty) such as that of the "self-model" is used as a foundation for the mind. In order to achieve scientific respectability, the notion is backed up by the whole repertoire of usual scientific-friendly epistemic entities: "information-processing", "emergence", "operation", and "self-model". The notion of "self-model" is proposed as the cornerstone on which the mental manifold develops. However, it is a move that shifts the weight of the mental domain on the intermediate notion of model. It does not solve it.

We are not against the notion of the model. It is a very useful notion like that of information. However, if one speaks of modelling as a generic way to say that, because of certain modifications in the neural structure, a brain is able to deal in a causally efficacious way with certain events, fine. The problem is that one cannot use the word "model" as if it would refer to something real.

One may see a wooden structure of Santa Maria del Fiore in Florence and may conventionally say that it is a model. Likewise, nobody is going to see a model of the world in someone's brain. The use of the word "model" is purely instrumental. The wooden model is a physical entity that is used as a model of a church. There is no physical test to ascertain whether the wooden structure is either a model of a church or a wooden structure that, by chance, resembles some existing, past or future religious building. Furthermore, the fact of being considered as a model does not determine any difference in the causal behavior of the wooden structure. The wooden structure is a physical entity in its own respect. It has a shape, a color, and a mass. It occupies space, and it reflects light. The model is an instrumental role we attribute to a physical structure.

Likewise, in my brain, there is nothing in addition to my neurons whenever they interact causally with the external world. If one uses the term "model" to refer metaphorically to the behaviors of neurons, it is fine. However, if one infers something out of the fact that the neurons model the external world, one may fall into the trap of one more ontological promissory note. The fact is that both with information and with modelling we have contracted ontological debts that we are never going to pay back. We have used terms that are supposed to support the ontological weight of the mind. Such terms are nothing but ontological promissory notes. All ontological payments are postponed in the future.

Ontological promissory notes lift the obligation to address the ontological status of consciousness. However, they do not solve the issue of the nature of consciousness. On the contrary, they entail a dualistic picture of the mind-body problem. In fact, such epistemic entities suggest the existence

of an ontologically light domain that is the place of the mind (see Fig. 1). It is a form of dualism because it suggests that only in the brain something special takes place (information processing, modelling, representation, and so forth). Further, it suggests that this special event (which is nothing but an epistemic entity) is the necessary (and perhaps sufficient condition) for the occurrence of the conscious mind.

The resulting dualism is based on the internalist assumption that the mind is the outcome of what is going on inside the CNS. Yet, so far, both empirically and conceptually, there is nothing like the mind inside the CNS because 1) the mind has properties that do not fit with the standard view of the physical world; 2) so far it has been empirically verified that there is nothing like the mind; 3) if neuroscience rejects dualism there cannot be anything but physical stuff. The situation is embarrassing, to say the least. As we mentioned, the three more popular options are eliminativism, emergence, or dualism in disguise. Eliminativism is usually rejected because it appears to throw away the baby with the bathwater. Emergentism begs the question. Once something is emerged, is a new ontology needed? After all, emergentism is not an explanation but rather an admission of ignorance (Kim, 1999). Eventually, neuroscience considers the last option—namely, to resort to some vague entity that may seem to be acceptable from a scientific perspective and yet sufficiently vague both to promise a future explanation and to realize now some epistemic work.

Consider information again. If it were physical, it would add something to the physical system that realizes it. Clearly, it doesn't. As we have seen, since information is not physical, at the end of the day, two options are conceivable. Either information is a pure epistemic concept (i.e. a metaphor to speak of something else)—and thus it may not be used to support any physical phenomenon. Or, we may be serious when we refer to the ontology of information, but in this last case, we would opt for a dualistic picture of reality (Chalmers 1996; Tononi 2004, 2008). The hypothesis is tantamount to assuming the standard physical world *and*, on top of it, a level of information floating above. This is full-fledged dualism, and it brings with itself the usual bunch of issues such as interactionism, ontological prodigality, and truth-conditions. At least, this informational dualism would have the merit of being coherent. On the contrary, in neuroscience, many scholars seem to assume a fuzzy and vague intermediate view—although information is not a physical thing (that would be naïve), it is nevertheless real and thus it may be the basis for further phenomena (guess what? The mind). If this is not an ontological promissory note, we no longer know what it is.

Finally, we would like to spend a few words on the least favorite epistemic entity of choice—namely, the issue of representation. In neuroscience and cognitive science, it is a recurring mantra that a representation of X doesn't need to share the properties of X to represent X. It is a rather

obvious consideration as to conventional representations. Consider a traditional praise of the notion of representation:

> We must distinguish features of representings from the features of representeds . . . someone can shout "softly, on tiptoe" at the top of his lungs, there are gigantic pictures of microscopic objects, and oil paintings of artists making charcoal sketches . . . To suppose otherwise is to confusedly superimpose two different spaces: the representing space and the represented space (Dennett and Kinsbourne 1992: 149).

Every student in cognitive science and philosophy of mind learned that. And yet, when it comes to mental representations, many doubts arise. If

Table 5.1 A comparison between current discussion about the mind in neuroscience and traditional dualism

	Physical domain	**Intermediate vague notions**	**Mental domain**
Descartes	Extended substance		Thinking substance
Melzack (Melzack, Israel, Lacroix and Schultz 1997; Melzack 2001)	Neural impulses	Pattern and neurosignatures	Pain
Haynes (Haynes and Rees 2006; Haynes 2009)	Brain activity	Mental activity	Mental states
Haggard (Haggard 2002, 2004)	Neural activity	Coding mechanism Integrated representation	Conscious awareness
Kendrick/Naselaris/ Gallant (Kay, Naselaris, Prenger and Gallant. 2008; Naselaris, Stansbury and Gallant 2012; Nishimoto et al. 2011)	Brain activity	Mental content	Visual experience
Noë (Noë 2009)	Sensori-motor contingencies	Knowledge of sensori-motor contingencies	Qualia
Tononi (Tononi 2004, 2008)	Causation	Information integration	Qualia

your representation does not share any property with what it represents, how can it represent it? The traditional notion of representation seems to work conditioned to the existence of a conscious subject doing the dirty work—namely connecting through the intentionality/semantics/consciousness of the vehicle of representation with the content. The fact is that no one knows how to naturalize representations. Neuroscience did not succeed so far. Nor did anybody else.

4. CONCLUSION

The main reason why neuroscience resorts to various forms of dualism in disguise is twofold: first, the assumption that the physical basis of the conscious mind has to be internal to the CNS; second, the empirical/conceptual failure to find such a basis inside the physical scope encircled by the CNS. It is an empirical failure because so far there is no evidence as to the occurrence of anything with intentionality, unity, quality, first-person perspective, and duration. It is a conceptual failure because the ontological foundations of the received standard physical domain have been defined so to exclude the aforementioned properties.

The result is that many authors consider an intermediate level made of epistemic entities that do not have any ontological weight such as information, computation, modelling, representation, symbol manipulation, and so forth. This level, which is nothing but an ontological promissory note, has the same role once assigned to the mental substance without apparently committing to the same metaphysics. This is an unfortunate strategy since consciousness is a real phenomenon and as such, it must spring out of real phenomena. In the lack of any unexpected empirical breakthrough, the current dualism in disguise does not have the resources to get anywhere. As a result, there are only two available options: either internalism is rejected or dualism is reconsidered in a more explicit ontological framework.

In the end, there are two possible options—either neuroscience rejects internalism and considers an ontologically revised physicalism, or it accepts explicitly its persisting covert dualism and brings into the open its dualistic framework.