Teleological markers:
Seven lines of hypotheses around Dennett's theory of habits

LUIS E. ECHARTE

University of Navarra, Pamplona
Institute for Culture and Society
Medical Education and Bioethics. School of Medicine
lecharte@unav.es

Abstract. Neuroscientists frequently use two folk psychology terms –self and consciousness– in formulating decision-making process models. According to Daniel Dennett, such notions lead them to dualistic view of brain functions and hence to dead ends or, at least, to less than optimal designs in the research field. Indeed, Dennett's position offers interesting insights that may help neuroscientists to comprehend the distinction between conscious and non-conscious behavior. In this paper, I first explore how habitual behavior could be defined from Dennett's model. Second, taking his view into account, I try to offer a better interpretation of habits. Particularly, I define habits as involving a teleological and preconscious process whose traits and dynamics are indistinguishable from computational ones, but this is so more in their ends than in their origins. Finally, I propose seven lines of experimental hypothesis to support this definition.

Keywords: habits; intentionality; intention; Intentional Stance Theory; consciousness; self; free will; voluntary control; agency; decision-making.

1. Two criteria of approach to habits

There is an ongoing debate in neuroscience about i) the relation between *habit behavior* and *conscious behavior* on the one hand, and about ii) the relation between *habit behavior* and *teleological behavior* on the other. For example, concerning the former, Ann M. Graybiel defines habits as automatically and virtually non-conscious behavior thanks to which the agent releases attention in order to "to focus elsewhere" (Graybiel 2008, 361). In this context, habits are a means toward optimizing conscious brain process. Some problematic questions emerge from this interpretation, such as what does virtual nonconsciousness really mean? Can we define it as unconsciousness? Is focusing attention always a conscious process?

However, most neuroscientists avoid the equivocal concept of consciousness in the categorization of habits and prefer to use the teleological criterion (ii). This is the case of Anthony Dickinson, who defines habits from a *mechanistic stimulus-response model*, where habits are unintentional and involuntary routines, that is to say, they are "simply triggered by the appropriate stimulus". For him, habits are the polar opposite of "purposive actions controlled by the current value of their goals" (Dickinson 1985, p. 67).

More recently, several authors have begun to locate habits in the teleological sphere. For example, according to Xavier Barandiaran and Ezequiel Di Paolo, Dickinson's definition of habits –as extremely rigid, repetitive sequences of behavior– respond to a trend imposed by cognitive sciences in which the behaviorist position has become dominant.

> In contemporary computational neuroscience and machine learning, habits are conceived as reactive behavior or probabilistic associations between stimulus categories (or measure of states in a problem domain) and action patterns (representing a set of motor commands that might alter the functional space of the problem environment) (Barandiaran and Di Paolo 2010).

For Barandiaran and Di Paolo, this approach suffers from a "lack of a clear theoretical building block for dynamical, embodied and situated cognitive

approaches". In contrast, they propose, "modeling habits as self-sustaining behavioral neurodynamic patterns where activity-dependent plasticity shows an extended temporal structure". Habit formation expresses, in this context, the spontaneous development of agents' preferences (Barandiaran and Di Paolo 2010).

Approach ii), for its part, is not free of problematic questions either. Some questions include: Is it possible to define goals excluding their phenomenological traits? And besides, does teleological behavior have any effect on higher-order deliberative behavioral control? In this paper I will study the strengths and weaknesses of Dennett's habit theory, which helps to reconcile approaches i) and ii).

2. Hidden dualistic assumptions

Dennett's first and strongest approach to habits is methodological. There are some "bad habits of thought" in modern civilized society that prevent us from developing a more consistent theory of consciousness (Dennett 1993, 252). For similar reasons, it can be said that these thinking patterns also make it hard to comprehend the very nature of habits –whether they are intellectual, emotional or behavioral habits. I will begin by explaining the contents of such trends of thought and their consequences.

Dualism is not only a matter of spiritualistic approaches. For Dennett, many self-proclaimed materialistic scientists accept implicitly large numbers of Cartesian assumptions. Throughout his work, Dennett examines five of the most widely spread and pernicious assumptions, including:

A. Consciousness is a mental representation.
B. Consciousness is an *epistemological machine.*
C. Consciousness is higher-order cognitive process.
D. Consciousness is the ultimate deliberative control.
E. Consciousness represents a privileged place to access one's own thoughts and decisions.

A blend of these five premises provides an overview of what Dennett calls the *Cartesian Theater* (Dennett 2013, 281).

Briefly, thesis A has to do with our inclination to believe that one of the main roles of neural system is drawing maps of inputs and outputs. Dennett takes this assumption as the source of all other dualistic assumptions.

From thesis A, it is easy to fall into Thesis B: accepting neural system as an information processor that primarily aims at taking the most accurate picture of outer and inner reality in order to facilitate an organism's adaptive behavior. In this sense, consciousness becomes the window through which we look into reality.

It is also because of A that we usually think C: the most accurate representation has to be created by the most complex neural process; the most sophisticated way of interpreting has to be sustained by the highest-order brain function (Dennett 1993, 187–193).

On the other hand, any representation, to be such, needs to be presented to someone in order to be known. Besides, we tend to think of consciousness as such a definitive eye because, through it, agents are able to control behavior (Thesis D). Indeed, knowledge and autonomy are two sides of the same coin in representationalism. There is no information without an agent; there is no agent (autonomous being) without information (Dennett 1993, 413–418).

Finally, A induces us to value introspection as an optimal method for observing the contents and causes of mental events. Thesis A pushes me to think that if I know what I am thinking and why I am thinking what I am thinking, it is because representations and observer are, in some way, separated. In other words, the observer controls all the inputs and outputs from a non-subjective point of view –Thesis E (Dennett 1993, 43–46).

Much of Dennett's work is devoted to showing how these five assumptions, as well as the way they interact, are very problematic. In doing so, he tries to undermine our predisposition to use them as axiomatic beliefs. In fact, this is the first step in understanding his approach—and to criticizing it. In order to do both, one can look to Dennett's objections in his analysis of Libet's famous experiment. Moreover, by focusing on this analysis, it is possible to study most parts of his theories on what habitual behaviors are.

3. Libet's observed

Benjamin Libet is mainly known for his experiments on unconscious electrical processes in the brain. They show how unconscious neural processes seem to precede some *volitional actions*, namely behaviors that are felt to be consciously and freely motivated by the subject (Libet et al. 1982; 1983a; 1983b; 1985). The problem is that prominent scientists, such as Michael Gazzaniga, Vilayanur Ramachandran and William Calvin, have interpreted this point as nearly definitive proof against free will.

For Dennett, there is a yawning gap between Libet's findings and the conclusion of these three interpreters. Nevertheless, the dualistic thinking habits of the latter encourage such a gap to go unnoticed. Concretely, thesis A, D and E form the bridge between both statements. Let's look at Dennett's argument:

> How can anybody write intentions in somebody else's mind? Don't we each have "privilege access" to our own decisions and choices? Not, not really. One of the major themes in Wegner's work is the demonstration, by a number of routes, that our knowledge of the relation between our thoughts and our actions (and between thoughts and other thoughts) has only the 'privilege' of ordinary familiarity. If I know better than you know what I am up to, it is only because I spend more time with myself than you do. But if you surreptitiously insert grounds for false belief into my stream of consciousness, you can make me think I am making 'free' decisions when it is you who controls my actions. The basic technique has been understood by magicians for centuries: Magicians now call it psychological forcing, and it is remarkably effective in able hands (Dennett 2004, 274).

Indeed, neuroimaging techniques are not necessary to prove that, on many occasions, there are a lot of unconscious factors that lend meaning to our thoughts or cause our intentions. Yet, there are many people who believe that information (inner and outer representations) belong to the physical world, while intentions belong to the conscious world. In other words, these people suffer from a nasty habit that prevents them from observing the simplest facts.

> It's very hard to change people's minds about something like consciousness, and I finally figured out the reason for that. The reason for that is that everybody's an expert on consciousness (Dennett 2010 [1969]).

Libet's experiment is useful for refuting thesis E, that is to say, to show why we are not experts on consciousness. For example, we do not always know, through an introspective act, when behavior is or is not intentional. However, this truth does not hamper the expression genuinely free acts, but intentions must not be reduced to the conscious world, or at least, to the phenomenical dimension of it. As we shall see, it is in this context that Dennett presents habit as a necessary part of voluntary control.

It must nevertheless be said that Libet himself tried to avoid a deterministic position, admitting the possibility of a veto power in consciousness. In his opinion, since our access to a representational map is thus delayed, the most we can do is intervene with last-minute triggers –in just a tenth of a second (Libet 2003). However, this is a very precarious idea of freedom. As Dennett claims, "this suggest that our conscious minds may not have free will, but rather 'free won't'!" (Dennett 2004, 230–231). Even worse, Libet does not exclude the possibility that decisions to veto (control) are based on unconscious processes that precede the veto (Libet 1999, 52–53).

However, talking about Libet's veto power, Dennett writes, "I certainly want more free will than that" (Dennett 2004, 231). In fact, with a new reinterpretation of the conscious-unconscious binomial, Dennett tries to overcome representationalism and all its dualistic mistakes, and thanks to that, to offer a stronger notion of voluntary control than Libet's one. In the next section, I will introduce the first pillar that supports his proposal: intentional behavior might be out of the agent's sight.

4. Voluntary control extended in space and time

Voluntary behavior is not always conscious behavior. Dennett agrees with the content of this statement, but not with the formula. Here, consciousness is used in an equivocal sense. For him, we should get used to identifying

consciousness with the first part of the sentence –voluntary control– and talk about the second part of the sentence using another expressions, for example, subjective impressions, performances of consciousness, qualia, etc. (Dennett 1993, 70–89). He thus gives primacy to the teleological criteria (approach ii) in the study of behavior. However, he understands the notion of goal very differently from Dickinson and Barandiaran. In such a definition, as I will show, subjective impressions (approach ii) are signs of its particular nature.

But let's not get ahead of ourselves. According to Dennett, the first step to comprehending intentions in a non-dualistic way is exposing the mistake of thesis D. To do that, we have to realize that all of the brain's chores are not gathered into one compact location "where everything could happen at once in one place –vision, hearing, decision making, simultaneity judging…" (Dennett 2004, 237–238). Voluntary control does not come at the end of the cognitive process, but rather, is integrated throughout it:

> Once you can see yourself from that perspective, you can dismiss the heretofore compelling concept of a mental activity that is unconsciously begun and then only later 'enters consciousness' (where you are eagerly waiting to get access to it) (Dennett 2004, 242).

For Dennett, human beings are *situation-action machines* most of the time —or at least until their behavior begins to bring unexpected consequences. In contrast, reflective moments, when agents prepare (in more than a tenth of a second) deliberative responses to daily life, are the exceptions. This preparation may also involve skill acquisition. In the latter, the cognitive, emotional and behavioral fruits of learning are progressively automated, which is to say that they are increasingly independent of rational triggers. Dennett values automation as one of the most evolved strategies of free will.

> Our lives are full of decisions to act when the time is ripe, revisable commitments to policies and attitudes that will shape responses that must be delivered too swiftly to be reflectively considered in the heat of action. We are the

authors and executors of these policies, even though they are compiled from parts we can only indirectly monitor and control. The fact that we can play ensemble music, for instance, show that our brains are capable of multitasking on a highly convoluted timescale, and it is all deliberate, controlled, and intended (Dennett 2004, 239).

Automatic and non-automatic behaviors are both voluntary processes that differ temporarily in their origin. The agent shaped the former long ago, while the second emerges in the very next moment after the making a decision. Therefore, for the latter, deliberation also triggers the action. This perspective makes it easier to explain why voluntary automatic behavior might be confused with non-teleological behavior. Agents have intentions towards goals but this does not mean that agents always remember their goals. Conversely, we can make the mistake of ascribing goals *a posteriori* to movements with non-teleological origins. Colloquially, we call this fakes *rationalization*. Again, thesis E does not correspond to reality since introspective criteria are not enough to confirm or deny voluntary control.

Dennett combats the Cartesian view that a goal is a state of consciousness or representation –a map seen by the big Eye. This is a widespread, but also counterintuitive, idea. In everyday language we often use expressions in which non-intelligent objects (mainly tools) are described as if they had purposes. For example, when the bell of an alarm clock begins to ring, we recognize a plan in its dynamics. Of course, this plan originates with the programmer, but this does not imply a continuous monitoring. In fact, an alarm clock's programmer is mainly characterized by being asleep. The persistence of the alarm clock's program is independent of the programmer's awareness. The important thing here is knowing to whom it concerns. The goal is executed by the alarm clock, but the programmer is the user of the goal's tool. Goals belong to their creators (or at least to their users) so, strictly speaking, they are programmer's goals. That's why it makes sense to classify an alarm clock's wakeup call of as intentional behavior. It is not the clock, but the programmer who ultimately wakes himself up into time and space.

This framework brings new dimensions to habits. It is true that they are helpful in decision-making in uncertain situations, as Barandiaran suggests (Barandiaran and Di Paolo 2010) and Libet's experiment confirms. Besides, Neuroscience has been long examining the role of habits in guiding behavior in low stress circumstances. Their effects in voluntary control, however, seem far more important, wider and mysterious: they shape our rational conscious preferences in daily life. In fact, Dennett's approach is useful for understanding why trying to uncover how habits make this third task remains one of the most important challenges in the neuroscientific study of decision-making. This I have called the alpha line of hypotheses.

5. Kinds of behavior

According to Dennett, an important difference exists between a computer (such as an alarm clock) and a brain in that the deliberative origin of many neural triggers is much less certain. Paradoxically, we are more convinced that behavior induced by internal organic means is more intentional –and human– than behavior induced by external machines. We already know which habits of thinking are partially responsible for such misunderstanding (Dennett 2001, 70–73). Yet, if introspection does not offer reliable criteria, how do we distinguish between intentional and unintentional behavior or, more concretely, between habits and non-goal directed behavior?

Given three big obstacles that stand in the way, in a logical analysis, it is difficult to objectify teleology once it has been created. First, goals do not depend on the complexity of the supposed agent. Continuing with the last example, alarm clocks and sand clocks express very different sorts of movements, but the existence of their goals are independent of them. In order words, teleology cannot be hastily defined as a higher-order cognitive process and herein, thesis C resounds again. Second, the attribution of one or several goals does not necessarily imply a physical modification of the tool. I can use my alarm clock as a paperweight or decorative object and neither use changes its dynamics. Third, goals do not require that they be accomplished, neither in the present nor in the future. A clock can be irre-

trievably lost or damaged and, at the same time, maintain its teleological descriptions, although its goal will never be achieved (Echarte 2008).

Dennett is aware of the problems linked to such a slippery concept. In order to solve it, he alludes to the Elizabeth Anscombe's definition of *acts*:

> What distinguishes actions which are intentional from those which are not? The answer that I shall suggest is that they are the actions which a certain sense of the question 'Why' is given application; the sense is of course that in which the answer, if positive, gives a reason for acting (Anscombe 1957, 9).

Like Dennett, Anscombe also denies that all intentional behavior must be under the agent's supervision. For example, it is for this reason that we can sometimes only know the real goals of our behavior after a certain period of self and external observation.

For Dennett, Anscombe gives us an optimal characterization of the class of intentional actions. "I can do no better than to give a *précis* of her analysis, making a few alterations along the way and then wedding the results to our emerging picture of awareness" (Dennett 2010 [1969], 184). Dennett's precision involves extending rational criteria to agents who lack central neural systems.

> – Why did you hit that man? I did it in order to steal his money.
> – Why did the clock strike one? The clock struck one in order to tell time.

We are more likely to accept the second sentence since the clock lacks a (complex) mind –thesis A and C. The clock does not know the time it reveals; time is not being *re-presented* to the clock. Again, Dennett urges us to think differently. What is the relation between mind and behavior? Classically, classifying behavior involves three kinds of criteria: spatial, phenomenical and deliberative. Thanks to them, we identify four kinds of causes in behavior. First, physical external causes –i.e., I fell because of the slippery snow. Second, physical internal causes –i.e., I fell because I suffered a heart attack. Third, psychical and irrational causes –i.e., I fell because I suffered a panic attack. And fourth, psychical rational causes

–I crouched down to pick up a dollar I saw on the floor. However, where do unconscious intentions fit in the above scheme? Indeed, according to Dennett, adjectives like physical and psychical create confusion in this categorization. For him, the best option is the simplest one: exclude phenomenical criteria. In other words, he identifies two kinds of causes to explain (teleological and non-teleological) behavior: externals/internals and rational/irrationals (Dennett 2013, 359–369). This reduction does not deny that the agent notices some internal and external causes, but that is independent of its causal power.

These considerations have practical projections in the neuroscience field. For example, it helps to conceptualize a hot topic in cognitive psychology: is consciousness necessary for conflict adaptation? To that question, Kobe Desender and Eva Van den Bussche answer that that it is commonly and mistakenly assumed in experimental sciences that "without consciousness we only be able to perform simple, reflexive behavior. Intuitively, we claim that we need consciousness to execute complex behavior based on our conscious intentions" (Desender and Van den Bussche 2012). In fact, Dickinson and Barandiaran also assume this perspective. The former assumes that habits are *non-controlled* and *absent-minded* behavior because he identifies them as lower-order cognitive processes (Dickinson 1985, 72). Similarly, in Barandiaran's model, there is a correspondence between the complexity of the brain process and the level of consciousness. The only difference is that, form him, habits are not just stereotypes, but rather are "plastic and self-sustaining pattern of behavior generated by emergent neurodynamic structures" (Barandiaran and Di Paolo 2010). Consequently they are only possible –originally– in the most evolved brains. In other words, habits have a conscious and controlled beginning, an unconscious life and, perhaps –thanks to the rehabituation process– some conscious interludes.

However, Kobe Desender and Eva Van den Bussche go on to say in the same paragraph, "An overwhelming body of research has shown that many cognitive processes can occur without consciousness meddling in" (Desender and Van den Bussche 2012). More importantly, many experiments confirm that unconscious stimuli are capable of influencing ongoing be-

havior (Kouider and Dehaene 2007). This recent data aims to confirm what Dennett was proposing back in the early 1980s: conscious content (like goals) can neither be defined as higher-order cognitive process (thesis C), nor as the ultimate deliberative control (thesis E).

The exclusion of subjectivity opens the way to explore new possibilities in classifying different sorts of behaviors:
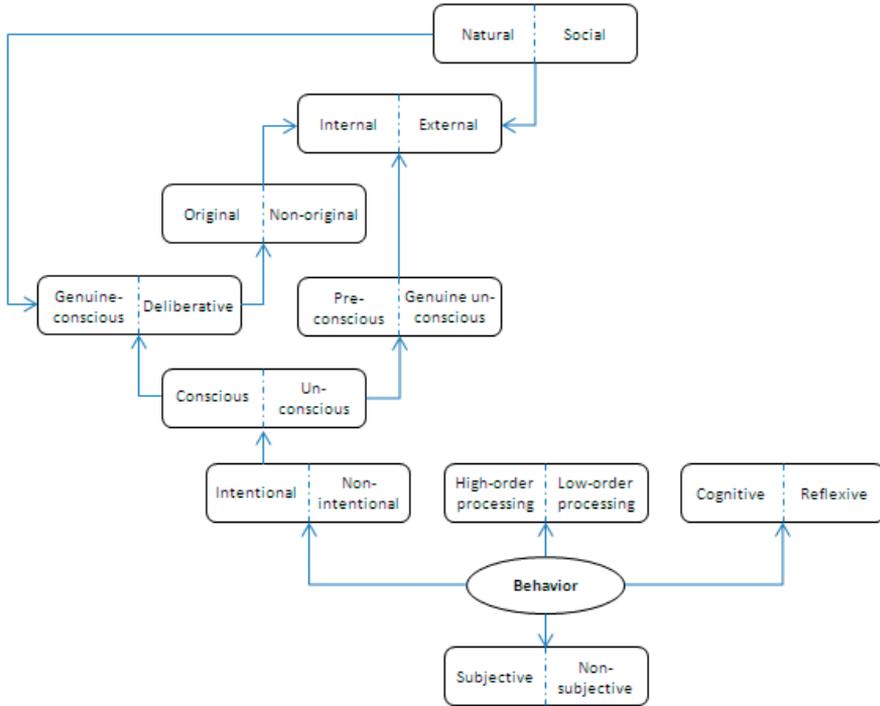
1) As non-intentional high-order processes

2) As intentional high-order processes

3) As non-intentional low-order processes

4) As intentional low-order processes

Besides, any of these four options, in turn, could be defined as either reflexive or cognitive. One specific reflex –classically, a nearly instantaneous movement in response to a stimulus– could require clever brain organization and a large amount of sophisticated algorithm computation. Conversely, one simple cognitive response –defined by Botvinick as fruit of the "system's ability to configure itself for the performance of specific tasks through appropriate adjustments in perceptual selection, response biasing, and the on-line maintenance of contextual information" (Botvinick et al. 2001, 624)– may not need complex information processing. It is precisely here where habits should be situated: behaviors whose origins are rational (intentional) (kind 2 or 4) but whose development becomes reversible reflexive behaviors (kind 1 or 3).

Lastly, any of these eight options could also be divided into behavior with or without introspective content. However, for the reasons that Dennett offers above, these sorts of subjective descriptions cannot be placed on the same causal level as the previous ones. Nevertheless, as we will see next, such phenomenological category is not sterile for understanding the deepest nature of our intentions and, by extension, habits.

Thus, all together, there are sixteen potential labels for describing behavior, where each one limits a portion of the broader-spectrum alternatives. As we will see in the following sections, we must add to this list another six more categories linked to intentional behavior (see Fig. 1). Another interesting neuroscientific tasks would be to detect (if they do exist)

Figure 1. Parallel descriptive categories of behavior based on Dennett's theory of consciousness

the neural correlates of the nine objective categories of behaviors. This I have called the *beta line of hypotheses* –the tenth category is subjective and therefore is beyond (but not irrelevant to) experimental methodology.

6. Computer mind

Dennett keeps his promise. Expanding voluntary acts toward the non-introspective world exponentially empowers agent's control. The price to be paid does not seem too high: to recognize that the development of intentional human behavior is similar to the development of some computation-

al processes. Of course, Dennett admits enormous quantitative differences, saying, the "brain can do lots of things that your laptop can't do" (Dennett 2013a, 110). However, this statement does not return to thesis C. Human central neural systems (as of now) not only involve a greater capacity to process data –i.e., complexity– but also do so in a novel way. For Dennett, this capacity is what differentiates consciousness from mind.

Minds are, according to Dennett, "just what our brains non-miraculously do, and the talents of our brains had to evolve like every other marvel of nature" (Dennett 2004, xi). Hence, computers have minds when they are able to imitate brain activity. In other words, computers are intelligence machines but not just any machine is intelligent. The following scene nicely illustrates the latter assertion. We could start with…

> mindless bits of mechanism, without a shred of mentality in them, and organize them into more competent mechanisms, which in turn could be organized into still more competent mechanisms, and so forth without apparent limit. What we might call the *sorta* operator is, in cognitive science, the parallel of Darwin's gradualism in evolutionary process… Both the humanoid robot and the hand calculator are made of the same basis, unthinking, unfeeling Turing-bricks, but as we compose them into larger, more competent structures at higher level, we eventually arrive at parts so (*sorta*) intelligent that they can be assembled into competences that deserve to be called comprehending. (Dennett 2013a, 96–97)

Against representationalism (thesis A), Dennett defends intelligence as not only the ability to accumulate and process information, but also as the ability to use it to survive. This position is not far from the way in which the Oxford philosopher Anthony Kenny explains knowledge:

> It is possible for a structure to contain information about a particular topic without having any knowledge about that topic. The train schedule contains the information about train departures; but it does not <u>know</u> what times the trains leave… A category difference is involved here. To contain information is to be in a certain state, while to know something is to possess a certain ability. (Kenny 1992, 108)

Concretely, for Kenny, knowledge is:

> the ability to modify one's behavior in indefinite ways relevant to the pursuit of one's goals... It is because the train schedule does not have any behavior to be modified by what is written on it that it does not know what the train departure times are. (Kenny 1992, 109)

In this last paragraph we find the major area of disagreement between Dennett and Kenny. According to the former, goals are not a product of mind, but are rather a product of consciousness. A system has a mind when it persists thanks to its capacity to extract information from the world, whether it likes it or not. The simplest homeostatic machine has a mind but it is not conscious. By contrast, Kenny claims that the acquisition of information is always a teleological phenomenon.

> I know of no better way of making the distinction between sense-perception and information-gathering that the one Aristotle used long ago in the De Anima, where he says that 'where there is sense-perception there is also both pain and pleasure'. (Kenny 1992, 111)

This discrepancy has significant impact on the debate about habits. However, before we dive into differences, let's walk-through the area of common ground: anti-representationalism.

Kenny, along with many other important 20th century philosophers (such as Ludwig Wittgenstein and Gilbert Ryle), criticizes *representationalism* as a theory of knowledge full of errors. For Kenny, the most serious and developed error is probably what he names the *homunculus fallacy*. Kenny, like Dennett, also thinks that this error has been present in neuroscience all along. In fact, several neuroscientists from the 1960s were already warning against it. The best example is Richard Langton Gregory, former Professor of Neuropsychology at the University of Bristol, who writes,

> The task of eye and brain is quite different from either a photographic or a television camera converting objects merely into images. There is a temptation, which must be avoided, to say that the eyes produce pictures in the brain.

> *A picture in the brain suggests the need of some kind of internal eye to see it -but this would need a further eye to see its picture... and so on in an endless regress of eyes and pictures* [Emphasis added]. This is absurd. What the eyes do is to feed the brain with information coded into neural activity –chains of electrical impulses– which by their code and the patterns of brain activity represent objects. We may take an analogy from written language: the letters and words on this page have certain meanings, to those who know the language. They affect the reader's brain appropriately, but they are not pictures. When we look at something, the pattern of neural activity represents the object and to the brain is the object. No internal picture is involved. (Gregory 1978, 9)

Anthony Kenny is even clearer when identifying the mistake, as in the following excerpt:

> In itself there is nothing misguided in speaking of images in the brain, if one means patterns in the brain, observable to a neurophysiologist that can be mapped on to features of the visible environment. What is misleading is to take these mappings as representations, to regard them as visible to the mind, and to say that seeing consists in the mind's perception of these images. The misleading aspect is that such an account pretends to explain seeing, but the explanation reproduces exactly the puzzling features it was supposed to explain. Sometimes cognitive scientists write as if the relation between mind and brain was that the mind made inferences from events in the brain. (Kenny 2007, 218)

Dennett and Kenny agree that brain creates internal and external maps. However, what makes a map truly a map is that it is used by the organism. Or better said, maps are first *maps-in-action* and, only later, can they be described as simple maps: organic states that are fruits of knowing and knowledge tools. Thus, both prefer to replace the words *representation* and *map* with *information*, which is less tied to the presence of spectators. But it must be noted that both also admit that this second concept suffers from the opposite problem: it substantiates the act of knowing, which amounts to extracting, to finding, information from reality. However, strictly speaking, there is no information at all before such act. *Knowledge* is knowing or fruit of knowing (Dennett 2004, 45–47; Kenny 2007, 108).

For example, when we say that an Indian tracker obtains information from the trail, it is not that there is information itself along the way. There are only stones, dust, foliage, etc., but not information. It is more apt to say that the trail is a knowable thing –useful for surviving or, at least (from Kenny's point of view), for hunting when acts of knowing are applied to it.

The picture changes when we talk about memory in the human brain and about database information. There, we say in a different sense that knowledge exists as potential and not as action. In fact, *information-in-action* is, in the database, the cause of *information-in-potency*, that is to say, a particular state of matter in a computational system– regardless of whether or not it is structural of functional. Dennett argues that the hardware/software distinction takes on its full sense in this context. It does not work for describing a trail, but it does identify some characteristics of computational and neural systems. On the hardware level, how physical mechanisms implement computational processes is described and, on the software level or the algorithmic level, it is "specified formally and rigorously the system's proper competence" (Dennett 1998, 75). It should be pointed out that software cannot always be identified with the act of knowing (mind). For example, we can install software packages in a computer with the sole purpose of improving order in our files. In this context, we are not using information but there are still programs running inside it.

In pursuing this line of thinking, it is possible to classify three informational states from reality without making any reference to subjective experiences:

a) Non-minded phenomena

b) Un-minded phenomena

c) Minded phenomena

Again, un-minded phenomena would be produced by the acts of knowing and have the capacity of recovering and enhancing such acts from the moment they are used. Besides, these three states could be assigned interchangeably to any of the non-intentional behavioral options described above: to high and low-order process, to reflexive or cognitive responses, and to behaviors with or without introspective content. Let us take an

extreme example. It is conceivable that a complex cognitive process may not help, or even be harmful, to an agent's adaptive strategies. This being the case, it should be categorized as non-minded phenomena. In parallel, it is not a big deal to admit that such auto-destructive computation is accompanied by a subjective experience or a non-minded conscious experience. Paradoxically, a non-minded phenomenon may have introspective contents.

Finally, within this framework, there exists what could be identified as a *proto-habit*: a minded or un-minded behavior whose origin is cognitive, but whose development becomes reversibly reflexive. For example, a computer can be programed to memorize the result of the most frequent long number multiplications. When one of these mathematical operations is identified then the computer is able to release algorithmic processes (and save time) by resorting to its memory functions. Of course, this is only a *sort of* habit because, for Dennett, habits are specific tools for the management of voluntary control. As I said, it would be interesting to find out if there are neural correlates of each one of the above defined categories. Now, it would be even more enlightening to know if brains use some kind of marker to distinguish among a), b), and c). I have called this line of research the *gamma line of hypotheses*.

7. The virtual machine

In the study of brain activity Dennett includes, in addition to hardware and software levels, a third level: the conscious level. Naturally, he is not *directly* referring to introspective mind content, but rather to human beings' most remarkable skill: voluntary control. Indeed, Dennett claims that a human being is free, but, at the same time, he argues against our habitual (and wrongheaded) way of thinking about it. Free will is indeed real, Dennett writes, "but just not quite what you probably thought it was" (Dennett 2004, 223).

In rejecting thesis E, Dennett comes to his most imaginative proposal: if human beings do not have privileged access to their own decisions and

choices, then how can we know and manage the tremendous number of mechanical influences on our behavior?

> These machines we inhabit simplify things for our benefit... In other words, we get a useful but distorted glimpse of what is going on in our brains... As creatures acquire more and more such behavioral options, however, their worlds become cluttered, and the virtue of tidiness can come to be 'appreciate' by natural selection. (Dennett 2004, 244, 247)

From this point of view, it is easy to understand why Dennett claims that consciousness is not necessarily a higher-order cognitive process (thesis C). Simplification is another biological strategy to survive, like the digestive function or language. Similarly, we are accustomed to saying that a screw is no more and no less complex than a shaft nut. Consciousness, just as other mental skills, might require complex operations, but it is possible to image basic ways of reducing a situation to particular components or relations, for example, by developing the simplest exclusion criteria before computing data or replacing, in a calculation, large numbers with short numbers. Of course, throughout the evolutionary process, only effective strategies have been preserved and thus we maintain the false illusion of thinking that its results are essential representations of the whole. Simplifications are not true, but they are useful. Again, consciousness is neither a mental representation (thesis A) nor an epistemological machine (thesis B). If we are looking for believable maps in the brain, those produced by the tool of consciousness should not be our first option. All these ideas are clearly expressed in the following quote, where Dennett compares the evolution of a brain with software engineers who try to make more user-friendly computers:

> Computers are fiendishly complex machines, most of the details of which are nauseatingly convoluted and, for most purposes, beneath notice. Computer-
> -users don't need information on the states of all the flip-flops, the actual location of their data on the disk, and so forth, so software designers created a series of simplifications –even benign distortions in many case– of the messy truth, cunningly crafted to mesh with, and enhance, the users' preexisting powers of perception and action. Click and drag, sound effects, and icons on

> desktops are the most obvious and famous of these, but anybody who cares to dig deeper will find a bounty of further metaphors that help make sense of what is going on inside, but always paying the cost of simplification. (Dennett 2004, 248–249)

For Dennett, the brain's centerpiece metaphor is the illusion of the Self (thesis D), because of which we assume that human beings are like a symphony orchestra directed by a conductor, permitting us to better survive in our environment. Surprisingly, Dennett is not a radical anti-Cartesian, arguing that Descartes is not at fault for everything. Instead, he claims, the *philosopher of doubt* merely gives us a theoretical basis that reinforces our predisposition towards taking on a dualistic way of being in the world.

Terms such as *free will* and *voluntary control* refer to a very specific kind of adaptive strategy: simplification. And yet, the main objective of adaptive strategies is not to change in first instance the behavior of the agent herself, but rather to change the behavior of other agents.

> In fact, we wouldn't exist –as Selves…– if it weren't for the evolution of social interactions requiring each human animal to create within itself a subsystem designed for interacting with others. (Dennett 2004)

Before *the self* there was *the other*. The term *you* is part of a strategy (simplification) to predict and modify other agents' behavior. For the latter, the generalization of this kind of strategy has caused the transformation of users themselves. Using Daniel Wegner's words, Dennett writes, "People become what they think they are, or what they find that others think they are, in a process of negotiation that snow-balls constantly" (Dennett 2004, 250).

Consciousness transforms organisms into a somewhat different machine through which they follow new rules and adopt new projects with regard to something whose existence is merely virtual. In this context, voluntary control is one of evolution's final achievements: by beginning to ask each other to do things, human beings are eventually able to ask things of themselves.

The notion of self introduces a new (virtual) dimension for describing complex behavior: intentionality. For Dennett, consciousness, in the tech-

nical sense, implies the ability to assign, manage and produce goal-directed behavior. But here, goals must be understood as the most simplified and tricky of concepts. They emerge in the context of the development of another particular brain skill, named by Dennett as the "advent of choice machines." They are...

> capable of evaluating the probable outcomes of candidate options prior to decisions. In the quest by brain to produce useful future, this is a major improvement over the risky business of blind trial and error. (Dennett 2004, 247–248)

Their predictions can be considered as proto-goals and responses caused by computation as proto-intentions. However, genuine goals amount to those that also contain a simplification-virtualization of said predictions. Dennett shows the difference between them by contrasting human intelligence against artificial intelligence:

> Suppose you are playing chess against a computer. You want to win, and the only good way of working toward that goal is to try to anticipate the computer's responses to your moves... How do you know what the computer would do? Have you looked inside? Have you studied its chess-playing program? Of course not. You don't have to. You make confident predictions on the quite obvious assumption that the computer 1. 'knows' the rules and 'knows how' to play chess, 2. 'wants' to win, and 3. Will 'see' these possibilities and opportunities for what they are, and act accordingly (that is, rationally) (Dennett 2013a, 77–78).

The adoption of these three premises to solve a problem is what Dennett calls the *intentional stance*: a strategy of interpreting an entity's behavior by treating it "as if it were a rational agent who governed its 'choice' of 'action' by a 'consideration' of its beliefs and 'desires' (Dennett 2013a). *Folk psychology* is another famous term used by Dennett to denote "the talent we all have for interpreting the people around us –and the animals and the robots and even the lowly thermostats– as agents" (Dennett 2013a, 73).

It is quite revealing that the first outlines of Dennett's Intentional Stance Theory appear in *Brainstorms: Philosophical Essays on Mind and Psychology*, a work that was published in 1981. Two years later, from neu-

roscience, Scott Fahlman, Geoffrey Hinton and Terry Sejnowski presented a framework in which brain essentially behaves as a *Helmholtz machine*, namely a machine that is able to model inputs in terms of probabilistic estimation (Fahlman, Hinton, and Sejnowski 1983). Since then, many neuroscientists have begun to observe how brain activity predicts future and correct predictions for improving such capacity (Dezfouli and Balleine 2013). In this regard, for many authors, the role of *striatum* –part of the basal ganglia– is to receive *predictive error signals* and integrate them with other kinds of data (Rangel et al. 2008). In light of these findings, Javier Bernácer and José Ignacio Murillo suggest that striatum are "the neural substrate of Andy Clark's proposal on practical action" (Bernácer and Murillo 2012, 418). In fact, following Dennett's line of thinking, Clark defends the *Theory of Bayesian Brain* where intentional actions are identified with *Action-Oriented Predictive Processes* (Clark 2013). We must acknowledge, however, that Clark seems to end up falling into some dualistic mistakes. On the one hand, as Dennett writes, Clark describes "the Bayesian brain as a 'hierarchical prediction machine'" (thesis C and D) (Dennett 2013c, 29). On the other hand, as Bernácer and Murillo point out, for Clark, cognition is "some sort of representation" (thesis A) (Bernácer and Murillo 2012).

There is still much to be learned on this topic, but Dennett's views help us to advance along the way. In order to understand the human brain, neuroscience is faced with the challenge of finding substrates for simplification functions –consciousness's essential skill. How are brain's predictions used, for example, to produce metaphorical concepts like *belief* or *intention*? This I have called the *delta line of hypotheses*. Among all formulated here, this last is, in my opinion, the most valuable tool for supporting or objecting Dennett's theory of consciousness.

8. I robot

Like mind, *consciousness* is also information-in-action. Now, the difference lies in the content of the information: Consciousness's content is related to folk psychological descriptions and narratives. Following this parallelism,

unconsciousness is the passive result of said activity, which can be reactivated without repeating the same processes that created them. Finally, *non-consciousness* may be defined as the lack of simplified data.

It should be noted that unconscious data are passive from the simplification functions' perspective, meaning that said information is not being used by the conscious system, but other mental processes may be operating with it. This is what happens, for example, when a weather website helps users to decide whether or not to take an umbrella when they go out. This sort of program makes recommendations for weekend plans without mind (the capacity to use the information for its own persistence) and without consciousness (the capacity to simplify its predictions). We can expect that something similar occurs in our brain. All intentional data and goal-directed behavior are always originated by the conscious system, but not all are consciously sustained and triggered. Finally, as in the case of the *states of mind*, all three of these *states of consciousness* have been described without making any reference to subjective experiences, but this does not imply that they lack introspective contents. Incidentally, by enumerating all the possible types of subjective contents (from non-mind to conscious intentional behavior), we understand how intricate the human stream of consciousness can be (in its classical, introspective sense), and also how confusing.

From there, we arrive to the main difference between teleological human behavior and the simplest teleological artificial behavior. Standard computers (so called non-sophisticated machines) may manage and execute goals like we do (replicating human behavior), but they do not create them. That is, by observing the origin of goals, we find the strongest difference between human and artificial teleological behavior: Human teleology is created by consciousness, whereas robotic teleology is produced by software programmers.

In this context, the question of to whom goals belong arises. Are computers with teleological behavior really agents? Common sense tells us that such machines are only tools for the achievement of human goals. But, following Dennett's line of thinking, this statement implies that, in some way,
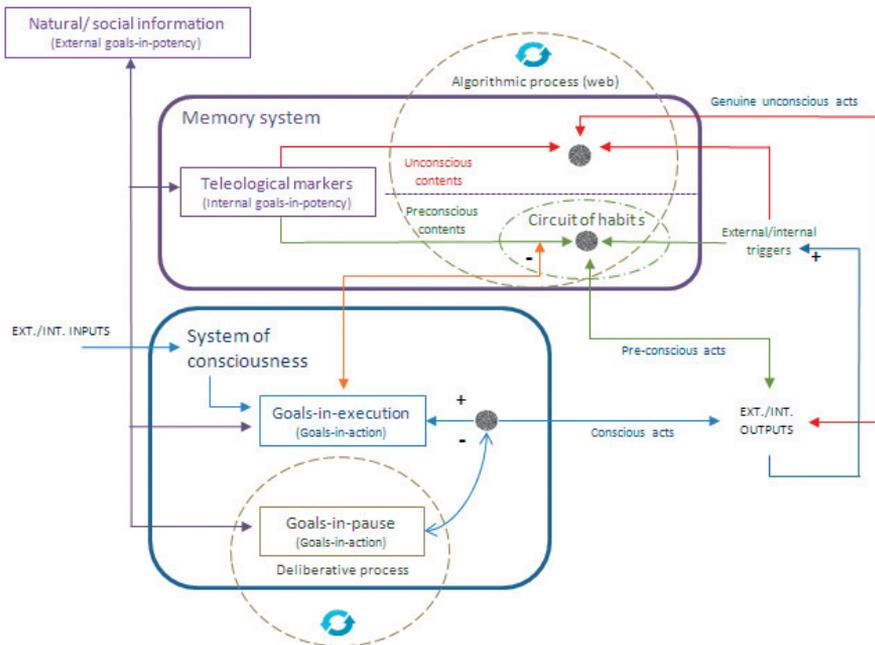
part of our mind is implemented in and executed by these tools. In the well-known paper *The Extended Mind*, Andy Clark and David Chalmers defend such a position, which can be summarized in two statements: a) cognitive processes can be constituted in part by elements that are not confined to the limits of the skin-skull barrier; and b) an agent's non-occurrent mental states (such as intentions and beliefs) can be partially constituted by realities that are not bound by the body (Clark and Chalmers 1998). When we reproduce or strengthen parts of our mind functions through mechanical gadgets we are not creating new minds, but rather extending our own. Therefore, the simplest computer's teleological behavior is not a *simulation* of human behavior, but actually human behavior. However, returning to Dennett's approach, the situation would be different if we were to design robots capable of predicting and simplifying inputs –i.e., replicates. Then, we could describe such sophisticated machines as authors of their own goals and, thus, conscious and free agents (Dennett 2013a, 168–174).

Clark and Chalmers do not think that a self-controlled robot is a perfect replication of a human agent. For this reason, they distinguish between *non-occurrent mental states* and *occurrent mental states*. The former depends on the agent's disposition to initiate the latter. For example, an occurrent belief is one that the agent is entertaining right now (Graham et al. 2007, 477–478). The difference between the two is similar to the difference mentioned above between *information-in-action* and *information-in-potency*, except for one element which amounts to dealing with, in Dennett's sense, conscious information. Nevertheless, according to Clark and Chalmers, it is impossible to guarantee that the phenomenal traits of any occurrent mental state are the same in humans and robots. However, for Dennett, this difference is not indispensable for defining conscious beliefs or voluntary behavior. Surely these robots will have their own subjective and particular experiences, but even if not, their behavior would still essentially be like human behavior (Dennett 2013a 341–346). Besides, we cannot be certain that *non-occurrent mental states* are exempt from subjective traits. *Beliefs-in-potency,* and many other functional dispositions, *may* exist, as stated, thanks to other complex brain dynamics.

9. Consciousness' on habits

Dennett's approach to human and robotic intelligence is directly linked to the problem of habits. However, before facing it, we have to introduce and confront two notions: *goal-in-execution* and *goal-in-pause* (see figure 2). The former refers to the origin of goals –those caused by predictive and simplification functions– while the latter refers to the fulfillment of the goal –usually directed toward motor behavior, but also toward emotional or cognitive changes. In this context, conscious (or voluntary) intentions could be defined as the behavior triggered by the very same goal formation process.

Figure 2. Circuit of habits based on Dennett's theory of consciousness

For example, *Martha asked Henry to keep her bag. He did*. The first step in Henry's behavior entailed (extremely fast) reasoning (predictions and simplifications) about the current situation, including mutual goals. Why did this occur? Henry can justify the execution of his goal even if he does not remember any previous deliberation. Following Anscombe's criteria, it is an intentional behavior. Following Dennett's criteria, it is also a conscious one. Henry is acting as a sophisticated machine.

Henry's instant response belongs to the most primitive and frequent type of conscious behavior. As Wittgenstein wrote in his *Philosophical Investigations*, "When I raise my arm I do not usually try to raise it (§622)." There is no previous deliberation process about what to do or, later, about the decision to do it. Wittgenstein express the same idea with a linguistic example: "One knows what one was going to say or wanted to say, and yet one does not read it off from some mental process which took place then and which one remembers" (§637). In fact, we do not usually prepare the thoughts in our mind before speaking. In all these cases, responses are caused by a type of a goal-in-act (an act of thinking or, as it was earlier referred to, *information-in-action*), what I call a *goal-in-execution*.

Let's image another scenario. *Henry was thinking for five minutes about whether or not to call Martha. She was very upset about his mistake. Finally, he called her*. Now Henry is indeed deliberating about his behavior and potential speech. Goals-in-act also go through Henry's head, but they are of another kind, what I call *goal-in-pause*.

John Searle defends a similar approach when he distinguishes between intention-in-action and prior intentions. The former is a goal of a non--premeditated voluntary behavior and the latter is a non-executed goal. In this sense, he writes: "Actions thus necessarily contain intentions in action, but are not necessarily caused by prior intentions" (Searle 1983, 107). However, I prefer to avoid both terms because they may lead to confusion about the passive nature of prior intentions, which are also acts of consciousness. Beyond this mere terminological controversy, it is interesting to point out how Elisabeth Pacherie uses Searle's distinction to defend two hypotheses that are very closely related to Dennett's theory about the social

origin of consciousness. First, prior intentions are quite similar to the state of mind formed by seeing others act and, second, primitive agent's prior intentions were derived from accidentally blocked or delayed intentions-in-action, which the agent progressively controlled and used (cf. Pacherie 1997). Again, the other first appears and, later, the self develops.

This leads us to question of what are unconscious (or involuntary) intentions. They amount to intentions for which there is no coincidence between the trigger's behavior and the origins of the goal's behavior. In other words, the cause of the action is not a goal-in-action, no matter whether executed or in-pause. Let's take an example: *Tom wanted to surprise his friend, Peter, and so hid behind the door. When Peter arrived home, Tom came out of nowhere and Peter punched Tom with a karate chop*. In this case, Peter's consciousness did not cause the unfortunate incident. Of course, it would not have been possible without Peter's earlier education and martial arts reflexes, where he learned to defend himself (namely, to acquire a goal-directed and articulated set of behaviors). Why did he hurt his friend? Peter can justify the execution of his goal: I thought I was being attacked so I used a karate chop. However, Peter's goal execution is not like Henry's goal execution. Martial arts training does not appear as necessary as it would have been if the situation truly required self-defense, as originally thought. In other words, Peter was not conscious about his reaction. Following Anscombe's criteria, it is still an intentional behavior. Following Dennett's criteria, it is an unconscious intentional behavior. In this instance, Peter acts as a non-sophisticated machine.

Using his introspective perspective alone, Peter would likely identify his behavior with Henry's behavior, namely, Peter would consider that the only difference lies in the speed of his thought process. We now know that this is only half of the answer: Henry's thought process was slower because he produced the original goal at the very beginning of its action. In contrast, Peter's behavior was caused by the convergence of the result of said actions– a *goal-in-potency*– with external or (non-conscious) internal inputs. I call these goals-in-potency *teleological markers* in that their execution is associated with the presence of triggers besides themselves.

Finally, it is not paradoxical to say that unconscious behavior may manifest some subjective traits, but, because of the speed of the initiation of such process (as I said, faster than the initiation of goals-in-execution), it usually goes unnoticed. Even more importantly, subjective experiences of conscious processes coincide in time with unconscious ones and some of them –especially those associated with goals-in-pause– last longer. However, there is a more important question that needs to be asked following this debate: Why does certain goals-in-potency become goals-in-action? Overall, why does certain information becomes conscious? According to Dennett, conscious is only a machine among machines (not the top one) so, there is no cause, only, natural selection:

> Mental contents become conscious not by entering some especial chamber in the brain, not by being transduced into some privileged and mysterious medium, but by winning the competitions against other mental contents for domination in the control of behaviour, and hence for achieving long-lasting effects... [The self] just is this organization of all the competitive activity between a host of competences that your body has developed. (Dennett 2004, 254)

The analysis of action showed here leads us to what I have called the epsilon line of hypotheses: comparing neural correlates of conscious behavior and unconscious behavior may reveal some clues about the functional and spatial (if any) limits between the system of consciousness and teleological markers' circuit (red and green lines in Fig. 2). However, choosing the right scenario for observing volunteers' brain activity is an important challenge. For that, as mentioned, all introspective criteria should be avoided. A final consideration nicely illustrates our wrong-headed predisposition to use introspective criteria, i.e., if Peter had been surprised by a real thief, it would be even more difficult to make him understand that his karate chop was an unconscious response.

And thus we turn to habits, which are a type of unconscious intentional behavior in which goals-in-potency can be quickly updated– becoming goals-in-action–, modified and executed as conscious acts (orange and blue lines). This process also partially breaks the automatic response triggered

by internal or external inputs (green lines), thus allowing agents tight control over unconscious behavior without cluttering the operations of the conscious system. How is a connection between sophisticated-conscious and non-sophisticated-unconscious machines possible? The mechanism that connects them is precisely related to the emission of *efference copies*, as Pacherie describes them:

> The idea is that when the motor centers send a motor command to the peripheral nervous system to produce movement, they also send a copy of this command to other centers that can in this way anticipate the effect of the motor command. (Pacherie 2012, 98)

The proposal herein amounts to the idea that habits emerge when the memory system (like a non-sophisticated machine) sends both a motor command to the peripheral area and an efference copy to the system of consciousness (the sophisticated machine), thanks to which, the self is able to anticipate this behavior, recognize it as caused by consciousness, and take control if necessary. On one hand, I call this kind of efference copy a teleological marker. On the other hand, following classic Freudian terminology, I call the resultant motor output a pre-conscious act. To conclude, it is important to note that the memory system does not always send efference copies to the conscious system or, at least, copies that allow the agent to rethink the behavior's original goals. Many could be the functional or accidental causes that explain such *non-rethinking*. Some of them are often referred to in nonspecific, common terms, such as forgetfulness, repression and illness. I refer to this last kind of intentional behavior as genuine unconscious.

10. Webs of beliefs

Let us return to the alpha line hypotheses. Usually, an isolated conscious goal (in execution or in pause) does not just depend on acts of prediction and simplification of external or internal phenomena, but also on other goals. As Donald Davidson wrote,

> [T]here is no assigning beliefs to a person one by one on the basis of his verbal behavior, his choices or other local signs, no matter how plain and evident, for we make sense of particular beliefs only as they cohere with other beliefs, with preferences, with intention, hopes, fears, expectation and the rest. (Davidson 1980, 221)

For instance, *Mark asked how much it was and then he paid for the movie tickets*. Here, paying is a conscious goal whose meaning content is determined not only by complex information about the meaning of money, a movie, etc., but also by many goals, such as "going to the cinema," "having a good time," "attending the movie" and so on. However, at the time of purchase, not all these goals and information were in Mark's consciousness. So, what and where were they?

Searle defines *background* as the set of all non-representational (non-conscious) abilities, practices and attitudes that are conditions of satisfaction for all intentional states. According to him,

> [The Background] is not a set of things nor a set of mysterious relations between ourselves and things, rather it is simply a set of skills, stances, preintentional assumptions and presuppositions, practices, and habits. (Searle 1983, 154)

In this quote, Searle gives the term *habit* a very specific meaning linked to motor behavior. It is worth pointing out that, a few years later, he wrote about another more classic sense:

> I think that much of Wittgenstein's later work is about what I call the Background. And if I understand him correctly, Pierre Bourdieu's important work on the "habitus" is about the same sort of phenomena that I call the Background. (Searle 1995, 132)

All the same, he uses the notion of background to define "certain category of neurophysiological causation" (Searle 1995, 129). Thus, in the context of the above framework, we can attribute such causes to what has been defined as teleological markers. Moreover, under such considerations, it

is possible to get a better picture of them: we should think of all these goals-in-potency like nodes in a web. An even better way of looking at them amounts to imaging such a web as a teleological algorithm where original goals-in-acts are adjusted before being executed. *Web of beliefs* is another name for this teleological matrix. In this context, a belief should be understood as a (cognitive, emotional or behavioral) state of mind that predisposes agents to think or to act in a particular way. Following this argument, part of our beliefs would be conscious (habits) and the other part would be unconscious.

Neuroscientists have only very recently shown an interest in in the study of the neural basis of human belief systems (Krueger and Grafman 2012). My ideas, if right about teleological markers, offer an interesting access to the essence of belief systems, facilitating the analysis of behavior patterns in pre-conscious and genuinely unconscious acts. These results might be combined with in-depth individual interviews (especially about deep-seated dispositions) and, finally, with neuroimaging. Lastly, findings about particular agent belief system(s) are a step on the way to elaborating an agent's core disposition profile, which might offer some predictive power in (not only unconscious, but also conscious) decision-making tasks.

11. Beyond brains

Another interesting part of background, according to Searle, is that it "can be causally sensitive to the specific forms of the constitutive rules of the institutions without actually containing any beliefs or desires or representations of those rules" (Searle 1995, 141). We already know that Dennett goes even further in claiming that the brain's social sensitivity is at the origin of consciousness:

> Plasticity makes learning possible, but it is all the better if somewhere out there in the environment there is something to learn that is already the product of a prior design process, so that each of us does not have to reinvent the wheel. Cultural evolution, and transmission of its products, is the second new

> medium of evolution, and it depends on phenotypic plasticity in much the same way phenotypic plasticity depends on genetic variation. We human beings have used our plasticity not just to learn, but to learn how to learn better, and then we've learned better how to learn better how to learn better, and so forth. We have also learned how to make the fruits of this learning available to novices. We somehow install an already invented and largely "debugged" system of habits in the partly unstructured brain. (Dennett 1993, p. 193)

In other words, social inputs are able to trigger an intentional response whose goal is not originated by the agent (never before been thought of), but rather is initiated by parents, governors, scientists, etc. So, first of all, we learn to do things and only later do we start to recognize goals as separate from their function. Unconscious acts predate conscious acts. Of course, this is not a new idea. See, for instance, the following quote from the *Nicomachean Ethics*:

> The virtues, then, come neither by nature nor against nature, but nature gives the capacity for acquiring them, and this is developed by training. Again, of all the things that come to us by nature we first acquire the potentiality and later exhibit the activity (this is plain in the case of the senses; for it was not by often seeing or often hearing that we got these senses, but on the contrary we had them before we used them, and did not come to have them by using them); but the virtues we get by first exercising them, as also happens in the case of the arts as well. For the things we have to learn before we can do them, we learn by doing them, e.g. men become builders by building and lyreplayers by playing the lyre; so too we become just by doing just acts, temperate by doing temperate acts, brave by doing brave acts. This is confirmed by what happens in states; for legislators make the citizens good by forming habits in them, and this is the wish of every legislator, and those who do not effect it miss their mark, and it is in this that a good constitution differs from a bad one. (1103a 25–1103b 6)

For Aristotle, habits are also a kind of learning in action. The Stagirite called this virtue, the most characteristic behavior of free agents, regardless of whether or not it has an internal or external conscious origin. In a similar way, Dennett claims that the acquisition of social habits (originated by others' consciousness) was, and still is, the first expression of freedom:

voluntary control extended in space and time, generation after generation. That is why human agents can neither understand the intentional content of many mental states with reference to themselves alone nor the relationship between themselves and their physical environment. Dennett refers to this kind of thought as *propositional attitude.* In this regard, he adopts an anti-reductionist position in suggesting that they cannot be entirely translated into neural descriptions (Dennett 1998, 124).

There are two senses in which the expression "extended voluntary control" can be understood. In the first sense, it refers to how a sophisticated-machine is able to colonize the mind of a non-sophisticated machine. For example, if we ask a little girl about her particular way of playing piano, she probably will not be able to respond; instead, her piano teacher will more likely be able to explain the reason behind the girl's technique. The consciousness of the teacher is being expanded toward the mind of this little pianist. It is not the girl who is acting –i.e., playing piano– but rather the teacher. In this context, teaching a baby is no different than using a calculator. The second sense has to do, according to Searle, with agents who acknowledge and participate in a social game. There, a social background is created thanks not to the sum of individual intentions, but rather to a single collective intentionality (Searle 1995, 126–127). Only in this second sense does it make sense to say that an agent is responsible for his or her unconscious socially originated behavior, which amounts to a shared responsibility. This is also the sense in which Dennett claims that voluntary actions are enabled not only by biological functions, but also by a kind of "software-sharing" (Dennett 1991, 194). He believes this to be a clear cognitive advantage. Human brains can take on novel routines of activity and execute novel algorithms without laborious conditioning thanks to the installation of "culturally evolved thinking tools – memes– that permit multiple layers of 'software,' creating indefinitely deep stacks of virtual machines that run on the underlying neural hardware" (Dennett 2013b).

With these considerations in mind, I want to finish this section with what I call the *zeta line hypotheses*: some pre-conscious and genuine unconscious acts are triggered by inputs that brain recognizes as social phe-

nomena. This is only possible if some teleological markers serve as detectors and interfaces among social inputs, computation of social information –external goals-in-potency– and agent's response. Internal goals-in-potency are differentiated by the way in which consciousness executes social preconscious behavior, i.e., by using the environment as a kind of memory that is external to the agent. Put in its simplest form, the brain would be able to identify and process some external inputs as fruits of (the other's) consciousness. Finally, coming back to Dennett's view, this social skill would not come as an improvement to internal memory, but rather quite the opposite. Social information came before neural information; metaphorically speaking, human beings learned to paint on cave walls before brains did.

12. Three problems in the science of consciousness

Dennett's ideas about the social phylogeny of consciousness shed light on some *categorial mistakes* that are present within cognitive sciences and neuroscience. The attempt to study consciousness is most frequently carried out through the design of non-sophisticated machines and through the observation of archaic living beings. For example, Barandiaran presents the field of evolutionary robotics (EA) as key to the richest notion of the organic decision-making process. There we find, according to the author, the root of different kinds of goal-directed behavior, including habits. For Barandiaran, spontaneous formation is at the origin of both. In the words of the author,

> Under different conditions, when no optimality criteria applies in terms of profitability, the history of interactions leads to the formation of particular preferences or habits on the agent. (Barandiaran and Di Paolo, 2010)

However, from Dennett's perspective, it is an oxymoron to talk about spontaneous preferences in a non-sophisticated robot, which would amount to speaking about the spontaneous emergence of a robot capable of printing legal money on the moon. On the contrary, only robots with predictive

and simplification functions are susceptible to being described with psychological concepts. Besides, the meaning of the term "preference" (like "teleology") refers to a kind of behavior among sophisticated machines. If we want to assign preferences to a sophisticated machine, we also need to recreate the same social matrix of intentions around it. In other words, we have to produce millions of similar machines interacting for millennia. Therefore, for practical reasons, the best method to study habits lies in creating a robot capable of integrating into human social life.

Yet again, from Dennett's point of view, habits, and the various capacities to store and use information in general, do not emerge automatically during just any part of the cognitive process, no matter the degree of complexity, but rather are additional and separable capacities (Dennett 2010 [1969], 51). Consciousness and habits should be thought of like language. Human beings did not begin to speak thanks to brain development, but rather because they learned how to influence others' behavior and to incorporate mechanisms to optimize such social strategies. Goals and parallel computing are kinds of tools in the same way that orange juicers and coffee machines are tools. An orange juicer simply cannot evolve from even the most sophisticated coffee machines.

Following the same line of argument, Dennett criticizes authors who try to ground the teleological ascription of living organisms in their homeostatic properties. This is also true of, for example, Matteo Mossio, who argues that "biological systems, as organizationally closed and differentiated self-maintaining systems, possess the necessary properties for adequately grounding both the teleological and normative dimensions of functions in their current organization" (Mossio et al. 2009, 822). According to Mossio, when an activity pertaining to the system becomes a necessary condition for the system itself, then it is possible to identify the goals of the system with a non-arbitrary criterion.

> To the question 'Why does X exist in that class of systems?' it is legitimate to answer 'Because it does Y.' In the case of the flame, for instance, the existence of the reactions of combustion is explained (at least in part) by taking into

account their consequences and, in particular, the generation of a macroscopic flow of energy (the flame), which in turn maintains the conditions in which the chemical reactions of combustion may occur. Similarly, organizational closure grounds normativity. Because of the organizational closure, the activity of the system has an intrinsic relevance for the system itself, to the extent that its very existence depends on the effects of its own activity. Such intrinsic relevance, we hold, generates a naturalized criterion to determine what the system is 'supposed' to do. In fact, the whole system (and its constitutive processes) 'must' behave in a specific way, otherwise it would cease to exist. Accordingly, the activity of the system becomes its own norm or, more precisely, the conditions of the existence of its constitutive processes and organization are the norms of its own activity. (Mossio et al. 2009, 825)

For Dennett, this kind of explanation extrapolates the concept of teleology to contexts where it loses its essential meaning. This is firstly because attempting to trace the tracks of teleology in the simplest phenomena is like trying to trace the tracks of currency exchange in a flame. For Dennett, the concept of goal is a specific survival tool –like antlers on a deer. Not every living being has antlers, nor does every living being possess teleological behavior. Assigning normative behavior to the simplest biological systems is, for Dennett, a mistake of anthropomorphization. He attributes it to the human congenital predisposition to apply social tools to any dynamic system. In fact, for him, there are "true believers that really have beliefs and desires" and non-true believers that "we may find it handy to treat as if they had beliefs and desires" (Dennett 1998, 22). Besides, Mossio assigns goals or relevance to a singular system appealing to its dynamics or inner elements. However, as earlier stated, this would be like saying that there is an intrinsic goal in the act of signing a check or claiming that the pen used to sign a check has intrinsic relevance in that act. On the contrary, we can only find and understand the goal of signing and the pen's relevance by observing the social rules followed by its user.

Dennett's objection can be strengthened with other complementary and classic arguments, typically known as the *naturalistic fallacy*. First, Mossio is confounding the normativity of ethics with the normativity of

logic. Goodness cannot be derived from necessity. The fact that X needs to do Y in order to exist does not necessarily imply that X *has to do* Y. That is only true if we admit that it is *good* for X to exist. But, why is existence good and nonexistence somehow bad? If we have difficulty thinking about the goodness of nonexistence it is, among other things, because of the *illusion of persistence*, as I call it. Most of our knowledge of reality refers to the ways in which inert and living beings persist. Besides, most of our moral decisions are wrapped up with the survival of someone or something. All this leads to a *mental habit* that makes existence automatically *appear* as a doubtless value (Echarte 2012). The champion of nonexistence makes no noise.

However, it is possible to think of the universe in a different way by, for example, paying attention to the apoptotic propriety of any system –i.e., by turning to all the process involved in its destruction. In this sense, Martin Heidegger wrote than any system is always a being toward any possibility –including its own destruction. Indeed, this last way of being-in-the world may give meaning to behavior. Systems do not need to be defined in terms of perdurability. This is precisely what Heidegger refers to, in more poetic words, when he claims that the *Nothing nothings* [in the original German, *nichtendes Nichts*]. The *nihilating Nothing* makes and updates itself and manifests itself in the Being, but such activity is unnoticed by our biased way of seeing things. *Being-toward-death* is the way of living, according to Heidegger, in which human beings understand the *nihilating* forces and, because of them, awaken to the subjectivity of the preference for survival. In other words, they move from the comfortable dark night of objective values to the bright day of normative inquiries.

> With the fundamental mood of anxiety (dread) we have arrived at that occurrence in human existence in which the nothing is revealed and from which it must be interrogated. (Heidegger 1928, 90)

The very existence of any system depends on the effects of its own activity as *necrophilous* or *biophilous*, using now Erich Fromm words. That is to say, it depends on the behavior of "those who love death and those who love

life" (Fromm 1964, 38). On the basis of a *necrophilous* system, we can re-formulate Mossio's statements saying that: X does not disappear because it does Y. So, Y has not an intrinsic relevance for the system's (way of) existence.

Finally, it is reasonable to think, from an evolutionist paradigm, which Mossio and Dennett assume, that, thanks to the study of the cosmic blind forces of nature, we can learn how man's good and the evil tendencies may have come about. For example, it becomes clear why human beings usual-ly think that life is better than death. However, as Thomas Huxley writes, "in itself [cosmic evolution], it is incompetent to furnish any better reason why what we call *good* is preferable to what we call evil than we had before" (Huxley 1893, 66). Therefore, if there are objective values (or non-arbitrary criterion to assign to norms), they should be sought in other places. Unlike Mossio and like Huxley, Dennett does not believe in such a (internal or ex-ternal) place but it does not imply that he claims that values are governed by individual opinion. Dennett defends a constructivist approach where values depend on society's moral rules. To explain himself he uses the example of *victory in chess*: this valuable goal is not grounded in spiritual entities or reduced to the movement of any game piece, but is rather grounded in the rules established by the players (Dennett 2013a, 385).

Another interesting conclusion derived from Dennett's view is linked to the inadequacy of neuroscience's reductionist methodology. Undoubtedly, for the study of some mental phenomena –including many higher-order brain functions– it is useful to start from the simplest and lowest neural interaction levels. However, other mental phenomena can only be under-stood in certain complex conditions. This seems to be the case of habit-ual behavior, whose functions can be better observed and understood in the context of real social interaction. However, it should be acknowledged that, although there is much to be gained from the *in vivo method*, there are also important obstacles. The biggest obstacle has to do with *group stud-ies*. How can we get a homogeneous sample of subjects and a reproducible scenario? The impossibility of controlling, in vivo, so many brain varia-bles should make us consider alternatives approaches. One of the most

promising alternatives, to my mind, comes to be so-called *ultra-cognitive* neuropsychology, in which data are obtained from single case studies and theories are mainly constituted by coherence criteria (McCloskey 1993; Caramazza and Coltheart 2006).

The final misunderstanding has to do with the interpretation of the brain as an epistemological machine (thesis B) and the practical consequences of Dennett's theory of consciousness. For him, many scientists still think of neural inputs as shadows of reality that the brain may rebuild. Consequently, they assume that maps created by the brain are useful for survival based on their verisimilitude. In other words, trying to understand neuronal maps represents the way towards existence and, at the same time, an adaptation strategy. Hence, when human being finds that a belief is false, we tend to deflate its adaptive value and change the associated behavior. In this dualistic context, what would be the consequences of knowing that the self and free will are not as we thought? Could Dennett's conclusions radically change our whole way of life? For example, neuropsychologist Steven Pinker and the neurophilosopher Patricia Churchland both hold this position in arguing that the social eradication of the concept of responsibility would be an important improvement in social life. Pinker expresses the ideal of the epistemological machine in the follow quote:

> Man will become better when you show him what he is like, wrote Chekhov, and so the new sciences of human nature can help lead the way to a realistic, biologically informed humanism. (Pinker 2002, xi)

Consequences are less serious when we interpret Dennett's conclusions in the context of his theory of consciousness. Dennett defines himself as a "sort of teleofunctionalist" (Dennett 1993, 460). Briefly, this means that "*truth*" and "reality" amount to additional folk psychological terms that human beings use to refer to their relationship with the environment.

> Every organism, whether a bacterium or a member of Homo sapiens, has a set of things in the world that matter to it and which it therefore needs to discriminate and anticipate as best it can. Philosophers call a list of things deemed to exist an ontology (from the Greek word for 'thing'...). (Dennett 2013a, 70)

Consciousness (and unconsciousness) is the basis of our reality. There, *forms of life* (using Wittgenstein expression) are imaged and rules are shared. So, consciousness is not true or false in the same way that a chess board does not win or lose a chess game. Understanding this statement does not imply a radical change to practices. On the contrary, it helps us to fulfill those raised in society.

13. Intention and intentionality

Dennett's objection to thesis B is one of the most controversial issues in his theory of consciousness. He claims that all statements (including scientific ones) are truth because of their hidden, lasting and deep utility. With reference to free will or self, objective knowledge is not what we thought. We do not know the intrinsic property of anything. Following David Hume, Dennett rejects even the existence of (real) causes. "The 'customary transition' in our minds is the source of our sense of causation, a quality of 'perception, nor of objects.'" Paradoxically, he resorts to evolution to explain why "mind's great propensity to spread itself on external objects" aids survival (Dennett 2013c). It seems strange, however, that our grammatical fiction (using another Wittgensteinian expression) was selected from mechanisms of natural selection like any other biological trait. Besides, human theoretical models have predictive power in part because they have been developed using rules of logic and coherence. It is as if our consciousness might mirror the way reality actually is.

Dennett tries to offer a theory of knowledge that keeps a safe distance from dualisms and, for that purpose, he derives intentionality –one of mind's distinguishing characteristics– from intentions or goal-directed behaviors. This approach represents a safer starting point because acts seem more accessible to observation and verification than with that mysterious "*aboutness,*" as Dennett also called it (Dennett 2013a, 61–64). In his proposal, intentionality is identified with meaning and talking about the meaning of something is, ultimately, finding its place in our complex world of shared interests. Thus, intentions come before meaning, namely

meaning appears when we focus on a node in the fabric of our intentions. Dennett illustrates the relationship between intentionality and intentions using the following example:

> Some naïve Americans seem to think that dollars, unlike euros and yen, have intrinsic value. The tourist in the cartoon asks, 'How much is that in real money' meaning how much is that in dollars. (Dennett 2013a, 299)

Asking about the meaning of a reality is like to asking about the value of money. There is no point in talking about something's meaning as an intrinsic property considered in isolation from its causes and effects. From here, a question naturally arises: what is left over if I start subtracting intentionality from intention? The result is an ever purer mind –a mind without consciousness, that is to say, a disposition without meaning. In keeping with this view, Dennett claims that intentions (contrary to intentionality) could be reduced to simpler and simpler less evolved forms until dispositional responses are indistinguishable from the movements of so called inert beings. We observe this stairway of agency not only in the evolution of species, but also in the organic constitution of any living being. This perspective is called homuncular functionalism and involves the decentralization of the body in which Cartesian homunculus is replaced with a committee.

> The proteins that are the workhorses inside your cells are amazingly competent and discriminating little robots... The neurons that do most of the transmission and switching and adjusting work in your brain are more versatile and competent robots... They form coalitions that compete and cooperate in larger structures... All these levels higher than the basic atomic building blocks exhibit a degree of agency. (Dennett 2013a, 88–89)

For Dennett, at higher levels, the sense of agency is stronger because the "jobs done" require considerable discernment and even decision-making. However, as Dennett himself describes, it is only a sense –a prediction and simplification of what is happening. For him, agency is a term proper to complex processes with adaptive value.

It is here where Dennett's greatest error emerges. By justifying intentions from evolutionary assumptions, he ends up denying intentions, i.e.,

the teleological dimension of behavior. If intentionality is based on intentions and intentions are only a kind of homeostatic phenomenon, then intentions end up being a merely "meaningful" way of labeling complex dynamics. Dennett falls into a recursive argument and the house of cards collapses: implicitly, he defends a sort of neo-behaviorism (neuro-behaviorism) in which intentionality is reduced to the agent's external (mainly social) and internal (mainly neural) movements.

In contrast, Kenny is correct in asserting that goals are a prerequisite for identifying mind's activity and not the other way around: no information exists if it isn't useful to anybody. It should be pointed out that Kenny uses the concept of goal in two senses: first, to describe when animals do something for the sake of another and, second, to describe when humans act for reasons, namely for "abstracted goals" or what I have called above goals-in-pause (Kenny 1992, 38–39). For Kenny, only animals with language have a mind because only they can give a reason for their actions. This, however, turns out to be a too narrow definition of mind. Franz Brentano's older definition offers some valuable insights. He characterizes mind as any sort of relational phenomena (intentionality): "In presentation something is presented, in judgment something is affirmed or denied, in love loved, in hate hated, in desire desired and so on" (Brentano 1874, 88–89). With this definition, a cat scratching on the door to get into the house has a mind. The cat's intention is an expression of intentionality, as well as of its emotions, memory, etc. In fact, almost all psychological states make reference to some content, even the simplest perception. For example, from the cat's point of view, different inputs appear linked to a singular experience, but in a differentiated manner. According to Brentano's perspective, this phenomenon is not a re-presentation, but rather 1) a co-presentation, that is to say, inputs are not primarily presented to the agent, but rather to each other by and in 2) the act of knowing. In reference to 1), philosopher Antonio Millan Puelles describes intentionality as a *consectary auto-presence*: it refers to "something essentially heterogeneous in the apprehension of an object but indissolubly linked to it" (Millán Puelles 1967). Subjectivity is not a performance of the act of knowing but

the aim of knowing. In reference to 2), philosopher Leonardo Polo writes, "Intellectual operation does not consist in 'going to' know [an object] because in doing so [object] has already been known" (Polo 2006, 75). Subjectivity is not an effect of the act of knowing but the act of knowing itself.

Dennett holds an extreme position in claiming that intentionality is derived from intentions. However, Brentano's idea does not seem completely suitable either: intention is not just another concrete expression of intentionality and the best that can be said of it is that any act of knowing (for example, perceiving a noise) is intrinsically intentional. That is, with the co-presentation of inputs, we also access the nature of their relations: causality –if material, formal, efficient and final. Hume and Dennett are right in claiming that we cannot deduce effects from experiencing causes. They err, however, in their conclusion that human intellection creates these notions. The alternative position is to assume that they are learned: any act of knowing implies intuitive access to some basic notions like entity, good, truth, etc. Philosopher Alejandro Llano calls these kinds of simple apprehensions or presences conceptual immediacy. In other words, sensible immediacy would be accompanied by this second immediacy (Llano 1999, 293–294). Apprehension of the heterogeneity of reality is an example of sensible immediacy. It is as hard to believe that heterogeneity is added by mind as it is to belief that causality is added by mind.

Knowing something (intentionality) amounts to adopting certain behavior (intention). The connection between intentionality and intention is remarkable because it is a promising way to defend realism without falling into representationalism. In addition, it shows us, using Merleau-Ponty's expression, that mind provides openness to the lifeworld (to the "Lebenswelt"). Acts of knowing reveal to living beings the causal dimension of reality, making them sensitive to laws and rules. In this context, mind and moral awareness are intimately related terms: knowing is synonymous with perceiving that something should be done. Here we find the deepest sense of terms like "potency," "disposition," "use of information" or "function." Finally, it is in this context where we can legitimately say that agents have privileged access to their own thoughts and actions.

Presences –sensible and conceptual immediacies– are, by definition, objectively baseless. Using another Wittgensteinian expression, we can show them, but we cannot talk about them. Indeed, knowledge about the existence of causes is a first-person *subjective certainty*. They make objective knowledge –and communication– possible, although they are beyond dialectical rationality. Dennett does not understand this last point and thus ends up supporting pragmatic relativism. He chooses *the wrong end of the ball of yarn*: it is not possible to explain subjectivity (first-person world) toward objectivity (third-person world) without missing objectivity.

These critiques are not made with the intention of refuting the whole model of habits here presented and, in fact, many parts of the model could be improved by making only a few changes. The most important one has to do with the origin of goals-in-execution. Dennett claims that consciousness creates them toward predictive and simplification functions, but we could redefine them as skills of detection: consciousness neither mediates between reality and agents, nor does it distort reality, but rather enriches our appreciation of it, bringing reality closer. I call genuine-conscious behavior to the behavior triggered by these kinds of natural goals. In contrast, deliberative behaviors (triggered by goals-in-pause) may be classified as original and non-original behaviors. The formers derive from the abstraction of the genuine-conscious behaviors' goals. The seconds derive from original goals which human beings are able to modify. In my opinion, Aristotle is talking about original and non-original behaviors when he wrote:

> It is absurd to suppose that purpose is not present because we do not observe the agent deliberating. Art does not deliberate. If the ship-building art were in the wood, it would produce the same results by nature. If, therefore, purpose is present in art, it is present also in nature. The best illustration is a doctor doctoring himself: nature is like that. It is plain then that nature is a cause, a cause that operates for a purpose. (*Physics*, II.8)

No matter which of these two versions, deliberative goals are typical phenomena of human acts of intelligence. Our consciousness is able to separate completely goals from its execution, namely to abstract the concept of

goal. Because of that, we may talk about goals, manage them (as concepts) and give reasons of our actions. The same could be said of the other kinds of *presences* –as identity, truth, beauty. All minds are open to the transcendentals properties of being but not any kind of mind think them in the same way. In contrast, computers are just informational machines. As Thomas of Aquinas wrote:

> Now whatever lacks intelligence cannot move towards an end, unless it be directed by some being endowed with knowledge and intelligence; as the arrow is shot to its mark by the archer. (*Summa Theologiae* I, q. 2, a. 3)

In fact, machines, can operate with the fruits of consciousness but they are not sensitive, at least not yet, to the heterogeneity of the reality and therefore to causal relations. In other words, they are not able of subjective acts.

In the same way, preconscious and genuinely unconscious acts could be originated and triggered not only by social, but also by natural information. This is an important point because it means that the background that makes sense of our beliefs, desires and the rest of our intentional stances are not only neuronal and social, but also natural. Human minds are ultimately connected to each other and to the mind of the universe. In this regard, I make my own the words of philosopher Mariano Artigas:

> Our universe exhibits rationality, information and creativity; that it makes possible the existence of human beings who are strictly rational and creative. (Artigas 2002)

I will finish this paper with my last line of hypothesis –I call it with the Greek letter *eta*–: some pre-conscious and genuine unconscious acts are triggered by inputs that brain recognizes as natural phenomena.

Conclusions

In this paper I have defended some Dennett's arguments as useful for understanding why neuroscience is swamped with dualistic mistakes. Second, I have explained why Dennett's alternative leads to serious inconsistencies.

Third, I have defined habit as a teleological and preconscious dynamics whose development (but not its origin) is indistinguishable from computational ones. Around these ideas, I have proposed seven lines of hypothesis:

α. Habits shape our rational conscious preferences in daily life.

β. There are nine objective categories of behavior.

γ. Brain uses teleological markers to distinguish between non-mind, un-mind, and minded phenomena. Besides:

   1) Agents have one or more sets of interconnected cognitive, affective and behavioral dispositions,

   2) It is possible to ascertain, through in-depth individual interviews, an agent's core dispositions profile. Neuroscience could help to characterizes functionally these webs –especially of deep-seated dispositions,

   3) Such webs have a significant causal role in decision-making tasks –even in those where the target seems neutral. Or, to put it another way, an agent's core disposition profile has predictive power.

δ. Brain transforms predictive data into metaphorical data.

ε. Comparing neural correlates of conscious behavior and unconscious behavior may reveal some clues about the functional and spatial (if any) limits between the system of consciousness and teleological markers' circuit (red and green lines in Fig. 2).

ζ. Some pre-conscious and genuine unconscious acts are triggered by inputs that brain recognizes as social phenomena.

η. It is the same than ζ but here the triggers are natural inputs.

My final conclusions deals with what I consider the main clue for facing the mind-body problem, which amounts to choosing *the right end of the ball of yarn,* thus avoiding many nonsense questions along the way to resolving it:

i)   Subjectivity is essential to the act of intelligent.

ii) Recognition of subjectivity is part of the foundation of Science.

iii) Causes cannot be reduced to sensible or physical events.

iv) Teleological behavior can be mechanistic and vice versa.

Acknowledgments

References

Anscombe, G.E.M. 2000 [1957]. *Intentions*. London: Harvard University Press.

Aristotle. 1893. *The Nichomachean Ethics*. Trans. F.H. Peters. 5th Edition. London: Kegan Paul, Trench, Truebner & Co.

Artigas, M. 2002. "The mind of the Universe. Understanding Science and Religion." In *Faith, Scholarship, and Culture in the 21st Century*, edited by A. Ramos and M.I. George, 113–125. Washington: The Catholic University of America Press. https://www.unav.es/cryf/themindofuniversenotredame.html.

Barandiaran, X., and E. Di Paolo. 2010. "Modeling sensorimotor habits with neuro-robotics. A Reappraisal of the habit concept in psychology." Accessed 13 Oct. 2010. Personal blog: http://xabierbarandiaran.wordpress.com/2010/10/13/philogeny-of-the-notion-of-habit/.

Bernácer, J., and J.I. Murillo. 2012. "An Incomplete Theory of the Mind." *Front Psychol* 3: 418. DOI: http://dx.doi.org/10.3389/fpsyg.2012.00418.

Botvinick, M.M., T.S. Braver, D.M. Barch, C.S. Carter, and J.D. Cohen. 2001. "Conflict Monitoring and Cognitive Control." *Psychol* 108: 624–652.

Caramazza, A., and M. Coltheart, M. 2006. "Cognitive neuropsychology twenty years on." *Cognitive Neuropsychology* 23 (1): 3–12.

Churchland, P. 2013. *Touching a Nerve: The Self as Brain*. New York: Norton & Company.

Clark, A. 2013. "Whatever next? Predictive brains, situated agents, and the future of cognitive science." *Behavioral and Brain Sciences* 36(3): 181–204. DOI: http://dx.doi.org/10.1017/s0140525x12000477.

Davidson, D. 1980. "Mental Events." In *Essays on Actions and Events*. Oxford: Oxford University Press.

Dennett, D. 1981. *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge Massachusetts: The MIT Press.

—. 1984. *The Elbow Room*. Oxford: Oxford University Press.

—. 1993. *Consciousness Explained*. Boston: Little, Brown.

—. 1998. *The Intentional Stance*. Cambridge, Massachusetts: MIT Press.

—. 2001. *Kinds of Minds. The Origins of Consciousness*. New York: Science Masters.

—. 2003. "The illusion of consciousness" TED talk. Transcription is available on follow link:

—. 2004. *Freedom Evolves*. London: Penguin Books.

—. 2010 [1969]. *Content and Consciousness*. New York: Routledge.

—. 2013a. *Intuition Pumps and Other Tools for Thinking*. New York: Norton & Company.

—. 2013b. "How thinking tools populate our brains and turn them into minds." Department of Psychology Keynote Lecture (2 March 2013). Tuft University, Boston, Massachusetts.

—. 2013c. "Expecting ourselves to expect: The Bayesian brain as a projector." *Behavioral and Brain Sciences* 36 (3): 29–30. DOI: http://dx.doi.org/10.1017/S0140525X12002208.

Desender, K, and E. Van den Bussche. 2012. "Is Consciousness Necessary for Conflict Adaptation? A State of the Art." *Front Hum Neurosci* 6: 3.

Dezfouli, A. and W.B. Balleine. 2013. "Action sequences and habits: Evidence that goal-directed and habitual action control are hierarchically organized." *PLoS Comput Biol* 9 (12): e1003364.

Dickinson, A. 1985. "Action and habits: The development of behavioural autonomy." *Philosophical Transactions of the Royal Society of London. Series B. Biologica Sciences* 308 (1135): 67–78.

Echarte, L.E. 2008. "Mind on Nature. The Problem of Unconscious Teleology." *Anuario Filosófico* XLI/2: 383–414.

—. 2012. "Función biológica y meta-biológica. ¿Qué significa mejorar la condición humana?" In *Selección de Actas del I Congreso de la Asociación Iberoamericana de Filosofía de la Biología (AIFBI)*, 418–424. Valencia: Publicacions de la Universitat de València.

Fahlman, S.E., G.E. Hinton, and T.J. Sejnowski. 1983. *Massively parallel architectures for A.I.: Netl, Thistle, and Boltzmann machines*. Proceedings of the National Conference on Artificial Intelligence. Washington D.C.

Fromm, E. 1964. *The Heart of Man: Its genius for good and evil*. New York: Harper & Row.

Graybiel, M. 2008. "Habits, rituals, and the evaluative brain." *Annual Review of Neuroscience* 31: 359–387.

Gregory, R.L. 1978. *Eye in the Brain. The psychology of seen*. New York: World University Library.

Heidegger, M. 1978 [1928]. *What is Metaphysics?* London: Routledge. https://www.ted.com/talks/dan_dennett_on_our_consciousness?language=en

Huxley, T.H. 1993 [1893]. "Evolution and Ethics." In *Evolutionary Ethics*, edited by M.H. Nitecki and D.V. Nitecki. Albany: State University of New York.

Kenny, A. 1992. *The Metaphysics of Mind*. Oxford: Oxford University Press.

—. 2007. *Philosophy in the Modern World. A new history of western philosophy. Vol. 4*. New York: Oxford University Press.

Kouider, S., and S. Dehaene. 2007. "Levels of processing during non-conscious perception: A critical review of visual masking." *Philos. Trans. R. Soc. Lond. B Biol. Sci* 362: 857–875.

Krueger, F., and J. Grafman, eds. 2012. *The Neural Basis of Human Belief Systems*. New York: Psychology Press.

Libet, B. 1985. "Unconscious cerebral initiative and the role of conscious will in voluntary action." *Behavioral and Brain Sciences* 8: 529–566. DOI: http://dx.doi.org/10.1017/S0140525X00044903.

—. 1999. "Do we have free will?" *Journal of Consciousness Studies* 6(8–9): 47–57. http://www.centenary.edu/attachments/philosophy/aizawa/courses/intros2009/libetjcs1999.pdf

—. 2003. "Can Conscious Experience Affect Brain Activity?" *Journal of Consciousness Studies* 10(12): 24–28.

—., E.W. Wright, and C.A. Gleason. 1982. "Readiness potentials preceding unrestricted spontaneous pre-planned voluntary acts." *Electroencephalography & Clinical Neurophysiology* 54: 322–325.

—., E.W. Wright, C.A. Gleason, and D.K. Pearl. 1983a. "Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiative of a freely voluntary action." *Brain* 106: 623–642.

—., E.W. Wright, and C.A. Gleason. 1983b. "Preparation or intention-to-act, in relation to pre-event potentials recorded at the vertex." *Electroencephalography & Clinical Neurophysiology* 56: 367–72.

Llano, A. 1999. *El enigma de la representación*. Madrid: Síntesis.

McCloskey, M. 1993. "Theory and evidence in cognitive neuropsychology: A 'radical' response to Robertson, Knight, Rafal, and Shimamura." *Journal of Experimental Psychology: Learning, Memory and Cognition* 19: 718–734.

Millán Puelles, A. 1967. *La estructura de la subjetividad*. Madrid: Rialp.

Mossio, M., C. Saborido, and A. Moreno. 2009. "An Organizational Account of Biological Functions." *Brit. J. Phil. Sci* 60: 813–841.

Pacherie, E. 1997. "Motor-images, self-consciousness, and autism." In *Autism as an Executive Disorder*, edited by J. Russell, 215–255. New York City: Oxford University Press.

—. 2012. "Action." In *The Cambridge Handbook of Cognitive Sciences*, edited by K. Frankish and W.M. Ramsey, 92–111. Cambridge: Cambridge University Press.

Pinker, S. 2002. *The Blank Slate. The Modern Denial of Human Nature*. New York: Viking Penguin.

Polo, L. 2006. *Curso de teoría del conocimiento I*. Pamplona: EUNSA.

Rangel, A., C. Camerer, and P. Read Montague. 2008. "A framework for studying the neurobiology of value-based decision making." *Nature Reviews Neuroscience* 9: 545–556.

Searle, J.R. 1983. *Intentionality*. Cambridge: Cambridge University Press.

—. 1995. *The Construction of Social Reality*. New York: Free Press.

Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Blackwell.