Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 31–32 (2014) 45–60

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

The functional role of free-will illusion in cognition: "The Bignetti Model"

Enrico Bignetti *

University of Parma, Via del Taglio, 43100 Parma, Italy

Received 6 August 2013; received in revised form 12 April 2014; accepted 16 April 2014
Available online 30 April 2014

Abstract

When performing a voluntary action the agent is firmly convinced that he has freely decided to perform it. This raises two questions: "Is this subjective perception of free will (FW) an illusion?" and "Does it serve a useful purpose?". The answers are tentatively given by "The "Bignetti Model" (TBM) as follows: (1) The so called "voluntary" action is decided and performed by the agent's unconscious mind (UM) by means of probabilistic responses to inner and outer stimuli; (2) After a slight delay, the agent becomes aware of the ongoing action through feedback signals (somatosensory, etc.) that are conveyed to the brain as a consequence of its performance. Thus, the agent's conscious mind (CM) always lags behind unconscious activity; (3) Owing to this delay, the CM cannot know the unconscious work that preceeds awareness, thus the CM erroneously believes it has freely decided the action. Though objectively false, this belief is subjectively perceived as true (FW illusion). It is so persistent and deep-rooted in the mind that the CM is unwilling to abandon it; (4) The FW illusion satisfies a psychological need to secure the arousal of the senses of agency (SoA) and of responsibility (SoR) of the action. Both SoA and SoR inevitably lead the CM to self-attribute reward or blame depending on action performance and outcome; (5) Both reward and blame are motivational incentives which foster learning and memory in the CM; the updating of knowledge will provide new information and the skill required for further action (restart from point 1).

Keywords: The Bignetti Model; Intentional action; Free will; Sense of agency; Sense of responsibility; Cognition

1. Introduction

The American philosopher John Searle believes that mind and body are not two different entities; that consciousness is an emergent property of the brain, and that consciousness is a series of qualitative states (Searle, 1997). With regard to the old philosophical question of duality and FW, Searle is astonished that the problem of duality has not yet been resolved, and thus asks himself why we find the conviction of our own FW so difficult to abandon. He writes: "The persistence of the traditional free will problem in philosophy seems to me something of a scandal". Nevertheless, many thinkers have studied this issue and many papers have been written, but it appears that little progress has been made. He questions: "Is there some conceptual problem we have simply ignored? Why is it that we have made so little progress compared with our philosophical ancestors?" He is not able to provide a philosophical solution to the question, and rather than adding further proposals, none of which would be convincing, he bypasses the obstacle by stating that "the philosophical mind–body problem seems to me not very difficult. However, the philosophical solution kicks the problem upstairs to neurobiology, where it leaves us with a very difficult neurobiological problem. How exactly does the brain do

* Tel.: +39 0521 032710, mobile: +39 342 8066908.
E-mail addresses: enrico.bignetti@unipr.it, biriko@icloud.com.

http://dx.doi.org/10.1016/j.cogsys.2014.04.001
1389-0417/

it, and how exactly are conscious states realised in the brain? What exactly are the neuronal processes that cause our conscious experience, and how exactly are these conscious experiences realised in brain structures?"

We agree with Searle when he claims to be astonished by this evidence, but we do not agree with him when he suggests that we should "kick the question upstairs to neurobiology" as if FW were not an intriguing issue anymore. This paper will attempt to take a significant step forward on this issue.

Material events can be described by an external observer as a chain of causes and effects which, in turn, may be causes for other effects and so on. Conversely, when we voluntarily cause an event, we do not feel that we are part of a chain; rather we consider our action to be the result of free will (FW). Wegner states that scientific explanations account for our decisions and the illusion of FW (Wegner, 2002). There must always be an objective mechanism, i.e., a precise relationship between causes and effects, underlying a voluntary action. We think that we consciously will what we are doing because we feel "free from causes" and because we experience this feeling many times a day (Wegner, 2002).

The obvious question is whether this deep-rooted subjective perception of FW is an end in itself or whether it plays some functional role in the voluntary action. In this paper, "The Bignetti Model" (TBM) suggests that FW (even if an illusion) is so deeply rooted in the agent's mind that it must be rooted in a real psychological mechanism of human cognition. The novelty of this model lies in its attempt to relate the psychological mechanism underlying subjective belief (illusion) in FW to the psychological motivation behind cognitive processes. The basic hypothesis behind TBM is that it is the sole idea of having FW that gives rise to the experiences of agency and responsibility of action. In turn, these experiences bring the conscious agent to judge the outcomes of the action and to rate the skill with which it is performed relative to his or her expectations.

2. Main actors in TBM

As an aid to the reader, here is a brief introduction to the main actors and their interrelationship.

2.1. FW and FW illusion

A popular definition of FW states that it is "an art for a particular sort of capacity for the rational agent to choose a course of action from among various alternatives" (O'Connor, 2013). Generally speaking a definition is worth since it is universally shared, i.e. all of us recognise ourselves in that definition. We believe that an outer observer of human behaviour like a machine or an electronic device could never come up with that definition since it cannot understand too many things of human mind, e.g. the meaning of "choice" or 'alternatives'. Then the definition could only be made through direct experience of the agent's condition, i.e. after choosing and performing an action. Under the belief of having freely chosen the action among all possible alternatives, the conscious agent perceives that FW is at work. Since the agent must be both the chooser and the witness (of him or herself), we need to clearly define the nature, limits, and subjective perceptions of the "rational" agent we are dealing with. For example, we must take into account that the idea of possessing FW is firmly rooted in the agent's psyche. Thus, the definition of the agent as "rational" seems limited since it necessarily excludes the agent's unconscious world.

Another issue arising from the definition is the suggestion that FW does not exist though we believe we possess it (FW illusion). We should ask ourselves if our will is really free since the action decision-making is conditioned by the prior stimulus and the best expectation of action outcome depends only on a cause-effect relationship. Being that our decision is always 'conditioned' we must logically conclude we are never free. Alternatively, there might be only one possibility to be really free and that is to decide an action by chance, for instance by throwing dice (eventuality which might be true of an insane mind). The paradox lies in the fact that a conscious agent believes in FW because he or she accepts the possibility that there might be conditioning even though he or she perceives him or herself as an agent who is "free from causes". Philosophy and psychology cannot mistake conditioning for a form of freedom so the question of why FW illusion is perceived by everybody needs to be resolved. A possible explanation is that FW illusion might simply serve as confirmation of one being alive and sane. Another possibility is that the illusion of FW might exert a functional role in cognitive processes.

These inferences may lend credibility to the theory put forward in TBM.

2.2. FW and consciousness

If you looked for a definition of 'consciousness' in a philosophical dictionary you would soon desist. The difficulty of providing a generally accepted definition is due to the gap that exists between the neurobiological mechanisms of brain and the apparently non-physicalist nature of the mind's activity (which keeps the debate on dualism going). There is general consensus that FW and consciousness are closely linked. In fact, the "freedom of will" (Van Gulick, 2011) has been thought to open a realm of possibilities, a sphere of options within which the conscious self might choose or act freely. At a minimum, consciousness might seem a necessary precondition for any such freedom or self-determination. How could one engage in the requisite sort of free choice, while remaining solely within the unconscious domain? How can one determine one's own will without being conscious of it and of the options one has to shape it?"

A brief survey of current thinking on the relationship between duality and FW is given to show that we do not

need a soul to subjectively perceive FW. In the western world belief in the existence of soul (immaterial essence of each individual, other than the body, source of consciousness, an agent having FW, responsible for thoughts and actions) depends to a large extent on the influence of religion and on one's social and cultural background. If men believe in God (whoever he may be) mankind's position is dominant with respect to the universe, but subordinate to God. Thus, the psychological weight of a subordinate position can be alleviated either by an irrational faith in God or by self-attributing a specific domain of responsibility with regard to material things (although this is still delegated by God). Conversely, if men do not believe in God, the individual self may be represented in different ways but cannot be identified with or considered the site of soul. In this case, duality becomes less relevant or disappears. Advances in neuroscience serve mainly to support the mind/brain identity hypothesis, showing the extent of the correlation between mental and physical-brain states.

Thus, there is a wide range of metaphysical positions in philosophy, as well as various theories of mind. Here is where we mention some of the more significant examples of contemporary authors who put forward very different theses on mind–body duality. The first are two philosophers and religious thinkers: Hans Jonas (1903–1993), and Emmanuel Levinas (1906–1995). Jonas proposed "Gnosticism" which concerns the dualism between two opposite or hierarchically dependent elements or forces, as in the case of matter (heavy, harmful and incompatible with mysticism and far from any spiritual realisation) and gnosis (elevated noetic or intuitive knowledge, the deep-rooted attitude of the soul to moral behaviour). Jonas defined Gnosticism as a "cardinal" dualism that governs the relationship between God and world, and correspondingly that of man and the world (Jonas, 1958). Levinas puts forward a philosophical perspective based on "the ethics of the Other" where FW employed exclusively for individual purposes would be nonsense. The Other cannot be made into an object of the self, and thus, cannot be acknowledged as an object. Levinas summed up his stance by saying that "Ethics precedes Ontology" (Ontology as the classic study of being). According to a famous statement: "The Other precisely reveals himself in his alterity not in a shock negating the I, but as the primordial phenomenon of gentleness" (Levinas, 1991). This is the moment in a person's life which requires self-responsibility towards "the Other," which is considered as irreducibly different. Levinas's obituary in *The New York Times* (Steinfels, 1995) read: "At the same time, the strict emphasis on ethical duty to 'the Other" as well as his commitment to Judaism, his resort to religious language and his many commentaries on passages from the Talmud and from the Bible separate Dr. Levinas from currents of post-modernism often viewed as radically skeptical or nihilistic. . ."

On the opposite philosophical front we find "Eliminative materialism" also called "Eliminativism". In this case, all or almost all classes of mental states (depending on the philosopher) are to some extent reducible to physical phenomena. Eliminativism relative to a class of entities means that class does not exist, so any form of materialism is eliminativist regarding the soul. Our mind is beset with false relics of common sense; moreover, any time we use words like intention, desire, love etc., in an attempt to describe the cause and the nature of events in our lives we are simply applying folk psychology. Neurosciences or other exact sciences will demonstrate sooner or later that such words or presumed states do not refer to anything material although real.

One of the most representative eliminativists, American cognitive scientist Daniel C. Dennett, eliminated duality from his cognitive theory by presenting the analogy of self as "a centre of narrative gravity". Thus, the self becomes not a physical entity but a purely abstract entity, a sort of folk-physics that is soberly known as phenomenology (Dennett, 1992a, 1992b). In a symposium held in Milan (Dennett 2001), Dennett began his talk explaining the magic of consciousness by stating: "It seems to many people that consciousness is a mystery, the most wonderful magic show imaginable, an unending series of special effects that defy explanation. I think they are mistaken, that consciousness is a physical, biological phenomenon – like metabolism or reproduction or self-repair – that is exquisitely ingenious in its operation, but not miraculous or even, in the end, mysterious. . .". This was the prologue presenting his basic approach (previously introduced in the famous book, "The mind's I" (Hofstadter & Dennett, 1981), which is completely unrelated to Cartesian mind–body dualism. The correct approach to his theory must be to remove the "subject". Thus, he states: "a good theory of consciousness should make a conscious mind look like an abandoned factory, full of humming machinery and nobody home to supervise it, or enjoy it, or witness it".

If Dennett denies the self, then two important questions need to be asked: (1) do we need a central agent (in charge) with direct responsibility for decision-making? (2) is the agent really or only apparently free when performing a purposeful action? Apparently, a deterministic perspective on life with a chain of causes which ultimately determine our actions is incompatible with belief in FW; on the other hand, indeterminism leaves no way for deciding a coherent action too. Thus, Dennett claims that the discussion 'determinism vs. indeterminism' is a red herring. At this point, Dennett suggests we reflect on the question: "Does neuroscience show that we don't have conscious free will?" He concludes by saying that several experiments have been conducted, but none are conclusive. On this basis, and taking into consideration that human frontiers are freely evolving in a Darwinian way, we will have to make some significant adjustments to our approach to FW and moral responsibility. So, if we go back into the mind of individuals we discover that "yes, we have a soul, but it is made of lots of tiny robots". There is no immaterial "soul" but the complex wiring and the teamwork of these robots that act as they are trained to; as they are governed, inspired,

adjusted and modulated by the cultural stuff entering our brain. This is a wonderful machine that manipulates 'memes' of information in an analogy with genes (Dennett, 2003). Dennett claims that in folk thinking if determinism is true then FW does not exist; therefore responsibility becomes a myth.

This raises the question whether in folk psychology, the complex system of robots in our brain can be deemed responsible for its actions in the way that a soul would be? If the answer is yes, then the robots in our mind could be held accountable by law. There are some pioneering experiments in which the participants in a task cheated a lot if they were previously convinced by reading a passage in a book that their brains are only a pack of neurons, that FW is only an illusion and that their choices are predetermined (Vohs & Schooler, 2008). In our opinion, those experiments seem to indicate that the agent's behaviour can be modified at any time, only if the idea of FW in memory contents is modified by external inputs. To this regard, TBM stands basically on the assumption that the meta-representation of self in a conscious agent (what we call self-awareness) stands on memory content, thus a transient modification of memory content may cause a very different representation of the self and of the inherent behaviour. A further assumption is that the conscious feeling of exercising FW in voluntary actions is fundamental to the self-attribution of agency and responsibility. Self-attribution of agency and responsibility poses Self (at least the meta-representation of it) at the centre of awareness waiting for the pronouncement of a blame or a prize, depending on the action outcome. This transient condition of the Self is a necessary prerequisite of human cognition.

2.3. FW and intentionality

In order to address the FW issue and its related questions, TBM must necessarily concern itself with conscious will and intentional actions. Intentionality can be defined as: "the power of minds to be about, to represent, or to stand for, things, properties and states of affairs" (Jacob, 2010). Therefore, we must consider TBM's agent to be of sound mind and dealing with reality, although we cannot claim with any certainty that either the motivations leading to the action or the critical evaluations of the outcome on the part of the agent might not cross over into conscious awareness.

We usually consider the purpose of acting as premeditated, i.e. as the mental causes of our actions only if we over-intellectualise. While we are consciously witnessing what we are doing, we are pervaded by the sensation of having wanted and caused it. Among a number of opportunities, we believe the action has been chosen freely, even though not always autonomously; this is the common sense that brings the idea of the existence of a FW. On the basis of these assumptions, two main models have been proposed in the current literature, a "Hard" and a "Soft" model, which support with different nuances the existence of FW (Gillett & McMillan, 2001). The former is a model that excludes any conditioning from interfering with the decisional action. From the rational point of view it is definitely unlikely: logic indeed asks us to consider the premeditation of a target as the necessary mind "conditioning" to formulate a choice. The second model, the "Soft" one, suggests that decision-making is a way to follow rules. The "Soft" model foresees a sort of determinism because of "the necessity of "rule-following"; though, to block the simplistic causal move required to ground deterministic thinking, the "intentionality of rule-following" might be considered as the result of a conscious, consentient and individual choice. As you can easily imagine, the epistemological root of this model is ambiguous and develops as a typical residual complex unsolved by religious faith. It exists in countless versions and it is in vogue especially among philosophers and scientists who have to bear a weighty social and cultural heritage, mostly derived from occidental monotheistic faiths.

2.4. Conscious mind (CM) and unconscious mind (UM)

The agent's conscious and unconscious mind (UM and CM, respectively) will be defined prior to introducing TBM's sequential events. The most appropriate definitions of UM and CM used in the model were found in the fundamentals of psychoanalysis. According to Freud the mind consists of three different levels: (1) the "conscious mind" (mental processing of everything that we are aware and we can rationally think of. It partially includes our memory, at least that part of memory content we can retrieve into the domain of awareness); (2) the "preconscious mind" (the ordinary memory, i.e. that part of the mind that can retrieve information from or pull them into consciousness, while we are not consciously aware of this activity at any given time); (3) the "unconscious mind" (a reservoir containing affects, urges, feelings, thoughts etc. which is therefore beyond conscious awareness. Our behaviour and experience are steadily influenced by the unconscious, even though we are unaware of these underlying influences. The unconscious is dynamic and is sealed off from the conscious mind by a force which he referred to as repression).

Therefore, the CM in TBM seems to fulfil the criteria of the Freudian "conscious mind", while TBM's UM, which is not directly involved in the subjective experience of intentional action and conscious will, would seem to match the characteristics of both the Freudian "preconscious" and "unconscious mind. However, a clear-cut distinction between CM and UM is not realistic (for instance FW illusion is not the product of a "rational" mind, rather it seems a psychological product of an emotional mind). Thus we cannot exclude *a priori* some kind of overlap between CM and UM. As we will see, all three levels could be extremely important in decision-making, action execution and cognition-raising.

In TBM the term 'conscious mind' could be sometimes replaced by Freudian "ego" to indicate "that portion of the human personality which is experienced as the 'self'

or 'I' and is in contact with the external world through perception" (Encyclopedia Britannica). However, we generally prefer to refer to Freud's earlier theory of the mind, the topographycal theory concerning with the unconscious, preconscious and conscious mind. Thus we prefer the term CM to focus the reader's attention on its distinct role though complementary with the UM in cognitive processes; moreover, although CM resides in the ego, not all the operations of the ego are conscious.

As a final comment we should underline the analogy between the roles of CM and UM in TBM and the roles of "explicit" and "implicit" minds, respectively, in the flow of the individual experience according to Dietrich's review (Dietrich, 2004). The mechanism by which knowledge shifts from an unconscious state to a conscious state is one of the most fundamental questions of cognitive science and lies at the heart of consciousness research. In brief the intriguing results here are that explicit mind (i.e. higher cognitive functions mainly supported by frontal and medio-temporal lobes) and implicit mind (i.e. skill-based knowledge mainly supported by basal ganglia) are two functionally distinct though interacting domains of mind. Thus, several steps occur before knowledge is fully accessible to consciousness.

Moreover, self-consciousness is a transitory meta-representation of the highest order of mind. In fact the frontal activity intervenes during executive attention, which is necessary to amplify the task at hand until it becomes the exclusive content of the working memory buffer; then it disappears quickly. The content of the explicit system is rule-based, verbalizable and tied to conscious awareness. So the flow is always under a critical analysis of the subject before being externalised. In contrast, the implicit system is devoted to experience-based and repetitive skill. The flow can be more complex though related to behavioural automatisms. Moreover, its content is not verbalizable and can only be conveyed through task performance and is inaccessible to conscious awareness. The main advantage of the implicit system is its efficiency. In contrast to the explicit mind, the implicit system does not seem to be 'capacity limited'.

Recent advances in cognitive neuroscience have begun to identify the brain circuits underlying the explicit system. Evidence that the working memory buffer of the dorsolateral prefrontal cortex holds the current content of consciousness, coupled with evidence that the executive attentional network of the region is the mechanism to select the content, suggests that the explicit system is critically dependent on prefrontal regions. Further evidence also suggests that medial temporal lobe structures are involved. All this leads one to infer that the explicit mind is evolutionarily more recent. This hypothesis is consistent with the view that information processing is hierarchically structured in animals with a highly developed prefrontal cortex. The functional hierarchy is devoted to exhibiting the most sophisticated knowledge representation and explicit mental abilities in the highest-order prefrontal cortex (Dietrich, 2003). Given that the explicit system is subserved by prefrontal regions, it follows that a flow experience must occur during a state of transient hypofrontality that can bring about the inhibition of the explicit system.

The neural correlates of the implicit system are not so clear. The basal ganglia are implicated in procedural memory (motor and cognitive skills), but contribute also to priming, conditioning, and habituation. Moreover, further central evidence is that optimal performance involving a real-time sensorymotor integration task is associated with maximal implicit mental ability of the tasks execution.

The neurobiological evidence reported in Dietrich's extensive review based on electrophysiological data seems to corroborate a reductionist view of CM and UM in TBM.

2.5. FW and the "Error Theory"

According to Wegner's point of view FW illusion is a subjective feeling that arises when the agent is convinced that he is doing an intentional action 'free from causes' and this feeling is reinforced many times a day. Thus, one may objectively argue that FW illusion is a by-product of the infinite repetition of a paradigm in which the subject is both the agent and the witness of the action. Conversely, a conscious agent can think about his FW as a genuine causal constituent of the action but he is just deceiving himself. Since the idea of possessing FW is a subjective feeling that lags behind the action, the definition of FW given above cannot hold.

Other situations in human behaviour have also been attributed to intrinsic, unavoidable psychological errors. These cases provided the philosophical bases for the formulation of the "error theory". Historically, this theory was introduced primarily to discuss the truth or falsity of moral rules. The principles on which "error theory" can stand, lead to the inference that knowledge requires truth. Thus, if there is no moral truth, there can be no moral knowledge and moral values are purely chimerical (Landau, 2010). The philosophy of "naturalism" sees moral judgments as true and obeying the laws of nature (Kurtz, 2003), while its opponents claim that moral statements are not reducible to natural terms (Landau, 2004). Answering this question would be beyond the scope of this paper; nevertheless, the fundamentals of "error theory" offer an interesting analogy to the debate concerning the objective-subjective perspective on FW illusion. As argued above, the conscious agent is unwillingly drugged into 'believing' in FW though this belief is objectively false. The question is whether this apparent contradiction leads to a deadlock or, rather, is the necessary preamble to something else. As we'll see in detail in TBM, FW illusion is perceived not to drive the intentional action, but simply to make the agent feel responsible for the action and to foster further cognitive processes. This second hypothesis avoids the pitfall of the soul-body duality by making subjectivity of primary importance in cognition; this is a noteworthy difference from other cognitive models.

2.6. FW and the sense of agency (SoA) and responsibility (SoR)

Our model stems from the hypothesis that it is simply because a conscious agent without FW would mean nothing in its own eyes that the subjective perspective of FW is so difficult to abandon. The denial of FW would be a sort of suicide. We must, therefore, consider two different points of view but arrive at a single conclusion. If we embrace a reductionist approach (the author's view), brain and mind are the same thing. Thus the persistence of duality and the belief in FW both reside in a psychological error: the agent's mind identifies the self with a body-independent entity (or soul) which, however is a product of mind. Thus an endless circuit of false attributions is activated without the objective approval of any outside observer. Instead in this dualism the mind is a "different thing" from the brain, living a life of its own, and need not be vindicated by the brain. According to a dualist tradition, intuition to an attentive mind is so easy and distinct that there is no doubt about what we comprehend and that we should search for truth by the light of nature. In nature, our ego might not be in the same space–time dimension as the brain and brain events (Krader, 2010), then self-identification of ego with soul can neither be proved nor disproved by brain activity.

In summary, according to the reductionist view, the conscious agent erroneously believes to possess FW; while according to the dualist perspective the existence of FW might be true. From whatever point of view we address the question, we can infer, firstly, that the persistence of the idea of a body-independent spiritual entity instantiated in our mind is imperishable, despite the fact that the body is physically deteriorating (the inner sensation that accompanies the sense of self is "sameness," an inferential activity instantiated in the prefrontal cortex (James, 1980; Van Den Berg, Vogel, Josic, & Ma, 2011); secondly, that this sense of self brings with it the idea of possessing FW. The first-person perspective on FW existence may be a subjective experience rooted in fundamental human needs, that's why it is a globally shared phenomenon despite its blend of theism and atheism.

Except for inner or outer constraints, such as moral rules or physical limitations posed by the environment, respectively, human beings, deep down, perceive the same degree of autonomy in both decision-making and action execution. This raises the question of the roles played by the "Sense of Agency" (SoA) and the "Sense of Ownership" (SoO). The SoA is commonly considered a constituent of the sense of self and the conscious agent refers to it as the feeling of being causally involved in an action (Gallagher, 2000). Recognising oneself as the cause of an action requires specific mechanisms in the brain to link inner intentions for voluntary action with a body- or brain-dependent execution. Both are inner feelings that we consider essential for consciously deciding, executing and controlling our actions. Neuroimaging studies have previously provided support for the existence of a discriminating system, which enables the subject to attribute to himself or not the responsibility for an action in the brain (Farrer & Frith, 2002; Mcguire, Silbersweig, & Frith, 1996; Ruby & Decety, 2001; Spence et al., 1997). The capacity to discriminate between a first-person or third-person action perspective is finely modulated by specific brain areas (Farrer et al., 2003). It has been demonstrated that a lesion causing spatial neglect interferes with self-recognition of the body in movement (Daprati, Sirigu, Pradat-Diehl, Franck, & Jeannerod, 2000).

As a scientist I would like to be more optimistic than Searle. For this reason we will propose a psychological model in which, false or not, the identification of CM with a "free-from-causes" entity and the idea of possessing FW are both conditions suitable to foster cognition (cf. Bignetti, 1994, 2001, 2003, 2004, 2010). In this paper the compatibility of the new model with current literature is then analysed.

3. The Bignetti Model

We usually consider the purpose of an intentional action as premeditated only if we evaluate it objectively and intellectualize it (third-person perspective) whereas the very instant we do something we are beset with the sensation of having "wanted" and caused it. We believe we have freely "chosen" the final action from among various options. This is the basic premise for the existence of FW (first-person perspective). "The Bignetti Model" describes the sequence of events of a voluntary action as having 5 stages:

(1) The so called "voluntary" action is decided and performed by the agent's unconscious mind (UM) by means of probabilistic responses to inner and outer stimuli.
(2) After a slight delay, the agent becomes aware of the ongoing action through feedback signals (somatosensory, etc.) that are conveyed to the brain as a consequence of its performance. Thus, the agent's conscious mind (CM) always lags behind unconscious activity.
(3) Owing to this delay, the CM cannot know the unconscious work that precedes awareness; thus the CM erroneously believes it has freely decided the action. Though objectively false, this belief is subjectively perceived as true (FW illusion). It is so persistent and deep-rooted in the mind that the CM is unwilling to abandon it.
(4) The FW illusion satisfies a psychological need to secure the arousal of the sense of agency (SoA) and of responsibility (SoR) of the action. Both SoA and SoR inevitably lead the CM to self-attribute reward or blame depending on action performance and outcome.
(5) Both reward and blame are motivational incentives that foster learning and memory in the CM; the updating of knowledge will provide new information and the skill required for further action (restart from point 1).

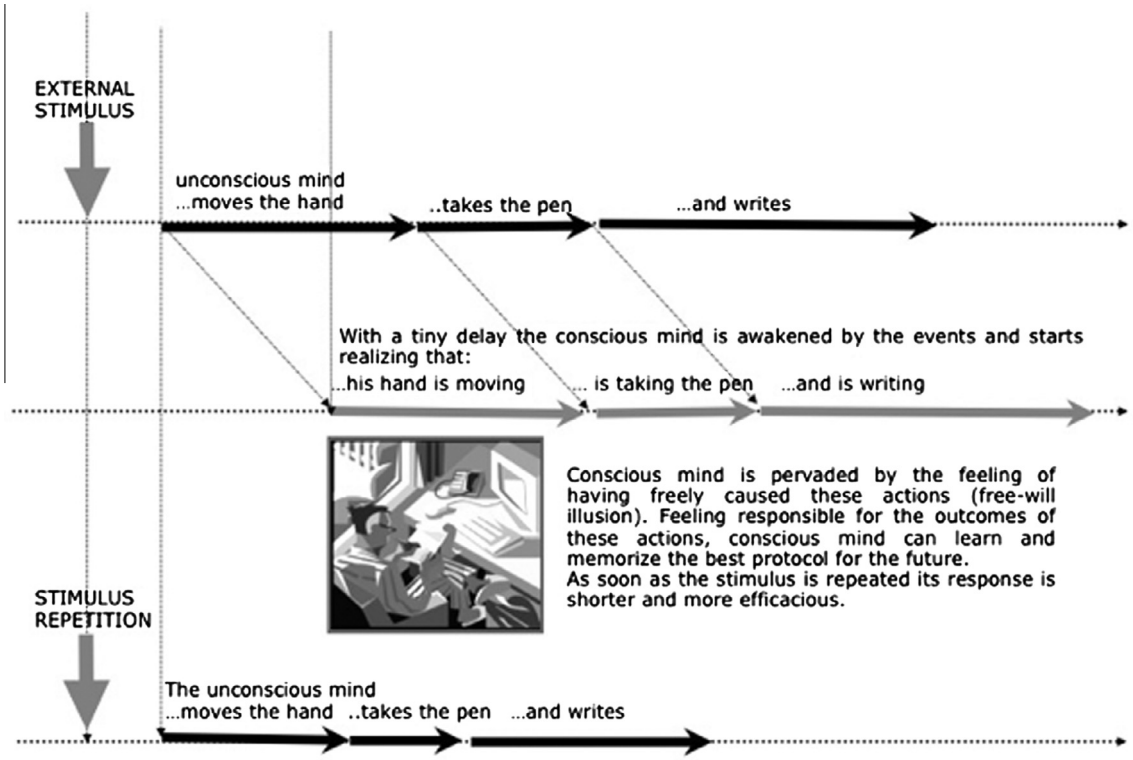

Fig. 1. Consciousness can learn but cannot decide. In this example the agent is sollicited to write down an answer in response to an external stimulus. Thus a series of actions are first performed by the unconscious mind (UM) (sequence at the top). With a tiny delay (times are only presumptive) the conscious mind (CM) is awakened by the sensory feedback signals. Thus, CM starts to witness each single action and, for each one, is immediately under the illusion of having freely and consciously caused it and, then by the sense of responsibility (SoR). The CM is then given to evaluate if the results match expectations and goals and, consequently, to self-attribute reward or blame; this constitutes motivational reinforcement to cognition (middle sequence). Finally, newly-acquired knowledge will be used by the UM to carry out other identical or similar actions (bottom sequence). An overview suggests that human knowledge evolves in a circular (or spiral) sequence of interventions from the UM to the CM and back to the UM.

A good example of a "voluntary" action whose timing can be analysed using TBM is reported in Fig. 1. On the basis of a probabilistic mechanism, the UM carries out its best actions in response to a stimulus. After a short delay, the feedback signals from this series of events awaken the CM in order to give it the opportunity to witness the events. Immediately, the CM is invested with the feeling of having decided them. Assuming that the performance of these actions is successful, the CM can verify, learn and memorise the correct protocol in response to the stimulus. In summary, our consciousness cannot decide an action but it can learn from its outcome and can update its memory store, thus providing the UM with the most accurate information possible in order to perform identical or similar actions in the future.

Several noticeable inferences can be drawn. First, TBM does not invoke the intervention of a soul or a body-independent entity to explain the sequence of events in an intentional action. The model is based on a psychological mechanism whereby every time it is awoken, the agent's CM erroneously feels as if it is a body-independent entity (or soul) and attributes to itself the role of a self-conscious causal agent, who decides and chooses "free from causes". TBM also claims that the idea of being a body-independent entity is instantiated in the agent's mind as a primary illusion, whereas the idea of possessing FW is only a by-product. Nevertheless, both illusions turn out to be an inseparable binomial apt for fostering cognition. The originality of this model lies in the causal role of FW illusion, not in driving the action but in fostering cognition. By means of this illusion the agent attributes to himself not only the role of player but also that of author and director in the 'film' of his life. By observing the overall sequence of events we may objectively propose in TBM that the subjective perspective of self and the concomitant FW illusion are tricks of the mind. As agent at the right moment he becomes aware of the ongoing action, he feels intrinsically dual. In conclusion, TBM reconciles the first- and third-person perspectives to give plausible roles of duality and FW in human cognition (Bignetti, 2013).

Unlike Searle we propose a self-consistent model in which we no longer need to kick the question of FW persistence 'upstairs to neurobiology'. The psychological and philosophical bases that account for the question have been posed. It is now neurobiology that should take it further. To this end, in a review dealing with the onset of a voluntary movement and the appreciation of whether this is voluntary or not, the author argues that FW is not the driving force behind it but is only the conscious awareness of it (Hallett, 2007). Since the sense of volition is a corollary response to motor discharges arising in the parietal lobe

and insular cortex, he concluded that FW was the result of introspection, subject to manipulation and illusion. The sense of agency must come from the appropriate match of volition and movement feedback, which is likely centred in the parietal area. The evidence presented and the argumentation in Hallett's work is of interest since it may possibly provide a neurobiological explanation of the first 4 points of TBM. The 5th point of our model, i.e. the proposal of a functional role of FW illusion in human cognition, should stimulate neurobiological research to further investigation.

The relative importance given to SoR which depends on FW illusion is effective in fostering learning and memory. Whereas the action execution is an obvious extension of inner intentions in response to specific stimuli, the "primum movens" of our knowledge, i.e. the link between action performance and conscious perception of causal agency, remains intriguing. From the discussion above, we may infer that a personal identity is psychologically installed in the agent's mind in order to observe autopoiesis: achieving the goal of self-organisation. In the overall picture, "Free Will" does not exist: it is only a belief of the inner observer. However, provided the inner observer survives, this illusion is justified since it is like an energy gear for such a cognitive system: it makes PI imagination work harder and better, i.e. it is the basic requirement for the reward circuitry operating at maximal efficiency; otherwise, according to Maturana and Varela (1980) and Varela, Thompson, and Rosch (1991) the system would disintegrate (Bignetti, 2001, 2003, 2004). Cognitive systems do not operate by representing world as a sum of independent components; knowledge is enacted as a series of distinct elements, inseparable from structures embodied by cognitive systems. On the one hand, the term "enaction" emphasises the growing conviction that cognition is not the representation of a pregiven world by a pregiven mind but is rather the enactment of a world on the basis of its history and the variety of actions that a being in the world performs; on the other hand, "embodiment" provides a systemic and dynamic framework for understanding how a cognitive self (a mind) can arise in an organism in the midst of its operational cycles of internal regulation and ongoing sensorimotor coupling. Another paper solely devoted to discussing the fundamentals of TBM in connection with the stimulating thought of Varela would be useful.

4. Point by point analysis of TBM

4.1. Point 1

TBM argues that 'free' decisions are determined by early brain activity. Libet's pioneering and controversial studies (Libet, 1983, 2004) on the timing of action decisions taken in the brain, observed the onset of early electrical activity, known as the "readiness potential" (RP), prior to the onset of conscious will. More recently, it has been shown that the outcome of a decision can be encoded in the brain activity of the prefrontal and parietal cortex up to 10 s before it enters our awareness. This delay presumably reflects the operation of a network of higher level control areas that begin to prepare an upcoming decision long before it enters our awareness (Soon, Brass, Heinze, & Haynes, 2008). This data is even more striking in the light of other research suggesting that the decision to move, and possibly the ability to halt that movement at the last second, may be the result of unconscious processing (Matsuhashi & Hallett, 2008). It is clear that mental processes that end with an intentional action pertain to a long chain of thought, each one being both the effect of the preceding event and the cause of the forthcoming one. Some preliminary studies of neuroimaging techniques, demonstrate that mind reading can anticipate an action by objectively interpreting the neuronal correlates with action intentions.

These studies are pertinent to our theory given that for information processing to take place both the UM and the CM share a sort of common neural 'language' or 'code' which is legible by brain circuits throughout the process described in TBM. Neuroimaging techniques are evolving to such an extent that the neural 'language' is also interpretable by a mind reader. A generally accepted view is that brain activity has evolved towards a probabilistic computation mechanism. Studies have shown (Koch, 1999) that each single functional component of a neuron, such as a voltage-gated Na^{+}-channel or an excitatory or inhibitory synaptic button, behaves in a stochastic way; however, if thousands of these neuronal components are engaged by stimuli from outside or from the network, their activity can be integrated, giving rise to a probabilistic (i.e. a statistically predictable) response. Thus, neuronal activity is predictable only if properly stimulated by the environment.

From a historical perspective, we have recently seen the advent of quantum mechanics, of chaotic non-linear systems, and of a renewed interest in the laws of probability; it is conceivable, therefore, that a dynamic model of brain function based on a statistic-probabilistic mechanism, e.g., the "integrate and fire" model (Lapique, 1952) may become the most popular. Brain activity based on a statistically predictable computation appears to fit natural events better than a pure stochastic or deterministic approach (Bullock, 1970; Deco, Rolls, & Romo, 2009; Koch, 1999; Lestienne, 2001). A turning point in research into the brain-mind relationship was the application of non-linear dynamics to neurosciences, which made the way for new brain activity models and the evolution of a mechanistic brain into a more dynamic system. To this regard, we will discuss two examples of probabilistic systems that could explain the agent's computational ability in TBM.

It is our view that the brain's intrinsic propensity for thought (a sort of compulsive "desire" to think) is a major dynamic propellant of the mind (Bignetti, 1994). Accordingly, the dynamic interaction of the brain with its surroundings of the "give and take" type was advanced by the theory of Continuous Reciprocal Causation (CRC)

(Clark, 1998). Years ago, a similar paradigm was deduced from the experiments of Ruch (1951): if one moves a finger forward to touch a small immobile target, the motion is not linear but involves a slight oscillatory movement towards the target, which becomes more pronounced in proximity to the target. This motion is the brain's spatial refining of the finger's approach to the target by means of trials and errors. This voluntary motion is also aided by involuntary brain areas and the feedback of vision and touch. Here, rather than a "give and take" mechanism, we should consider a "give, take and evaluate the transient outcome from action feedback" mechanism. The hand's position is relayed by feedback signals, step-by-step, so that the brain can perform a differential computation between the real and expected position. This brain activity is reasonably explained using Bayesian Decision Theory (BDT), which has been described by several authors (Kording & Wolpert, 2006; Norris, 2006; von Hofsten, 2004). BDT suggests that the computational brain behaves in a similar way to a probabilistic machine, in the sense that decisions are taken on the basis of statistical terms and functions which may become relevant to the decision; ambiguous decisions require larger statistical analyses. Subjective experience that fosters the acquisition of new knowledge may also be relevant for the fine-tuning of future decisions.

The CRC model appears inadequate in describing action-making unless we introduce a computational unit calculating the derivative of the position along the motion. It may not be necessary to upload or retrieve long or short-term memories; we know that sensory memory holds sensory information for a few seconds or less after an item is perceived (Atkinson & Shiffrin, 1968). This type of memory is outside cognitive control, and may last long enough for the trial-and-error paradigm to calculate and to adjust motion direction. Sensory feedback signals first awaken and then inform the CM of what UM has done with a slight delay. It follows that the theory that action encoding in sensory memories may last long enough to be conveyed to the CM, is also appealing to explain point 2.

In conclusion, we can say that TBM is compatible with the post-adaptive learning mechanism proposed by BDT. Long-term and short-term memories may also intervene to provide the unconscious and conscious mind respectively with useful information for action decision-making and the critical evaluation of action outcomes. The model is not in conflict with the computational probabilistic-deterministic ability of the brain which leads to predictable responses.

A second example concerns the "intelligent" behaviour of an oil droplet entering a water maze and finding the shortest way to the exit without making a mistake. The droplet behaves like laboratory mice after a long period of training (Lagzi, Soh, Wesson, Browne, & Grzybowski, 2010). This phenomenon is due to chemotaxis. The droplet and the exit of the maze are pre-treated with opposite ions so that the oil droplet is naturally 'pulled' towards the exit by the gradient. At least two conditions are necessary for this to happen (even without a brain): (1) a "pre-existing" knowledge of the goal and a deterministic self-attraction between opposite charges; (2) the probabilistic motion of the droplet that will favour it to cross the attraction field. There are many other examples of probabilistic behaviour in nature (enzymatic activity, microbial flagellar motion, etc.) and it stands to reason that if the effect exists before the cause of an action, the action is predictable. Using this analogy, when the effects achieved through intentional action are clear and unambiguous, the agent is consequently predisposed to accept and further interpret the incoming stimulus in a conditioned, non-free state, though perceiving an inner freedom from the causes.

An analogy may be drawn between these deductions and the hypothesis of "the brain's resting state" made by Northoff (2012). He retrieved Kant's hypothesis on specific intrinsic features of the mind that enabled the correct interpretation of the information delivered by an external stimulus. This ability of the mind may be dependent on the early onset of an intimate relationship between the mind and stimulus (readiness which may be described in operational terms as resting-state activity). Subsequent action is spontaneous and independent of the stimulus.

4.2. Point 2

The awakening of the agent's consciousness during action performance is made possible by at least two different mechanisms.

It has been known for more than a century that the brain generates its own electromagnetic field. This phenomenon is widely used in EEG, MEG and TMS. This, in conjunction with the evolution in field theories which were first introduced in Gestalt psychology, inspired McFadden who elaborated the "conscious electromagnetic field theory" (CEMI). As reported in several papers (McFadden, 2002a, 2002b; McFadden, 2006), CEMI is based on the idea that the combined firing of all the neurons in the brain generates a complex electromagnetic field which may induce a self-regulation of their activity. According to the theory, consciousness can be understood as an electromagnetic phenomenon produced by brain activity. The CEMI theory provides a realistic physical model that accounts for the subjective difference between conscious and unconscious mental processing.

McFadden (McFadden, 2006) examines several clues to nature and argues that the CEMI might provide a solution to all of them. For instance McFadden claims that we experience the influence of the CEMI field as FW. That is why willed actions feel so different from automatic actions: they are the effects of the CEMI field functioning as the inner cause. To this regard he argues that: " ...although like modern cognitive theory the CEMI theory views conscious will as a deterministic influence on our actions, unlike most cognitive theories it does at least provide a physically active role for will in driving our conscious actions...Our awareness (the global CEMI field) plays a causal role in

determining our conscious actions". By attributing a deterministic role in guiding purposeful actions to will, he claims the old Cartesian mind–body dualism has been resolved and a new matter-energy dualism has replaced it.

Though we appreciate McFadden's basic concepts, we do not agree with his view of the role of conscious will in voluntary action. In TBM, the agent has the perception of having FW but slightly delayed with respect to the true action (when his CM is awakened). We cannot, therefore, attribute to FW an effective role in deciding and executing an action. We can, however, attribute to the conscious agent a fundamental, psychological role in fostering learning and memory processes. Yet CEMI is an intriguing theory since learning and memory are cognitive processes that either require the presence of a conscious agent and occur only after the outcome of the action. Thus the awakening of consciousness in point 2 may well be explained by the reverberating effect of the electromagnetic loop as a consequence of the occurrence of the events in point 1.

The second point concerns the existence of Rizzolatti's "Mirror neurons" (Rizzolatti & Craighero, 2004). Mirror neurons could play a fundamental role in enabling the "self-mirroring" of action performance, allowing the agent to have direct experience of action outcomes. In Fig. 1, the self-mirroring effect could constitute the basic mechanism in facilitating the awakening of an inner witness prior to FW illusion.

In fact, mirror neurons represent groups of neurons that fire both when an animal is performing an action and when an animal observes the same action performed by another animal. These neurons have been observed in primates and other species, including birds. In humans, brain activity consistent with that of mirror neurons has been found in the premotor cortex, supplementary motor area, primary somatosensory cortex and inferior parietal cortex. According to Rizzolatti and colleagues, without action interpretation and imitation, social organisation and survival are impossible. Thus, we can assume that in humans there is a faculty that is dependent upon the observation of others' actions, known as imitation learning. Human cultural development could be based on this faculty.

The theory that mirror neurons can facilitate imitation has been emphasised and adopted by other groups. The neuroscientist Ramachandran demonstrated that mirror neuron activity was fundamental for a healthy mind, and believed that human evolution was mainly the result of imitation learning. This evolution was evidently Lamarckian because it was dependent on a horizontal spread of information through populations (Ramachandran, 2010). However, not all neuroscientists agree with Ramachandran's theory. One of the most plausible criticisms is that imitation requires recognition and recognition requires experience. Some researchers performed an experiment in which they compared motor acts that were first observed and then executed to motor acts that were first executed and then observed. The significant asymmetry observed between the two processes led these authors to conclude that mirror neurons do not exist in humans (Lingnau, Gesierich, & Caramazza, 2009). Despite criticism, the discovery of mirror neurons opened new routes in neuroscience that inspired other dynamic perspectives of brain function. A heuristic model, the "Shared Circuits Model" was introduced (Hurley, 2008), which suggested the existence of an intermediate system mediating a cognitive elaboration between incoming signals and intentional actions. Mirroring and the simulation of mirroring is one part of this artefactual dynamic system. Layered between the outer world and consciousness, this system enables human cognitive capacities for imitation, deliberation, mind reading, motor control and other functions via sensorimotor feedback.

Typical aspects of mind reading, such as the attribution of false beliefs to others, were demonstrated with 15-month-old infants (Onishi & Baillargeon, 2005). According to Gallese (Gallese, 2007) these results suggest that social skills dependent on these brain mechanisms develop very early, well before the development of language. There is a 'structuring' computational circuit within the premotor system that can operate in two ways. In the first, the circuit can organise action execution and/or action perception and imagination via neural connections to motor effectors and/or to other sensory cortical areas. In the second, the same system applies both to master language organisation and to yield 'abstract inferences'. According to this hypothesis the same circuitry that controls how to move our body, enables the understanding of the action of others and can, in principle, also structure language and abstract thought.

In this regard, it would be interesting to know if individuals are fully aware when "inner speech" is activated, in accordance with Baars (1998). This mechanism allows an individual to communicate and learn in order to adapt his actions to the environment for a homeostatic purpose (Maturana & Varela, 1980). On performing an action, we may not be aware of it but we can subjectively experience it by interrupting it and by putting ourselves in a meditative mood (Bignetti, 2004). The same occurs with the "inner speech" echo that somehow evokes an interior perception described by others (Edelman & Tononi, 2000), which probably corresponds to: "being conscious of being conscious".

4.3. Point 3

As soon as feedback sensory stimuli of the ongoing action are conveyed to the brain, the action's course becomes explicit to CM in a step-by-step manner (see the section above: "Conscious mind (CM) and unconscious mind (UM)" and Dietrich, 2003). Lagging behind UM, CM cannot see earlier UM's work; thus the agent believes it has freely decided the action. This illusion triggers a functionally useful sense of responsibility (SoR) in CM which exerts a positive effect on cognition (points 4 and 5), despite the fact it is based on an unavoidable psychological error!

Other aspects of human behaviour have also been attributed to intrinsic and unavoidable psychological errors.

These cases provide the philosophical bases for the formulation of "error theory". Historically, this theory was first introduced to discuss the truth or falsity of moral rules. A forerunner of the theory was Hume. According to Hume, no moral knowledge stands on pure truth, and thus it is purely chimerical (Hume, 1739). This issue is still controversial. The philosophy of "naturalism" sees moral judgments as true and obeying the laws of nature (Kurtz, 2003), while its opponents claim that moral statements are not reducible to natural terms (Landau, 2004). This author affirms that 'good' can neither be described in terms of pleasure and pain nor can it be the product of any of the natural sciences (physics, biology). Moral nihilists consider morality to be constructed, i.e., it is a complex set of rules and recommendations that may provide a psychological, social, or economic advantage for its adherents, but is otherwise without universal or even relative truth in any sense (Landau, 2010). Thus, ethical theories resemble genetic material that can naturally evolve and benefit from spontaneous mutations (favourable errors), which increase the organism's skilfulness in a society.

In this regard, TBM proposes a human cognitive mechanistic theory selected on the basis of FW illusion. This subjective perspective stands on a psychologically unavoidable error, but it ultimately leads to cognition. Objectively, we can consider the individual perspective of CM as an error, however, what is epistemically objective may not necessarily be ontologically objective. Searle (1997) argued that "where consciousness is concerned, the appearance is the reality", thus the truth can be discovered and evaluated by any interested party if we find the way to reconcile 1st-order and 3rd-order perspectives.

4.4. Point 4

Bodovitz thinks that we become aware only when some of the content of our underlying cognition is changed by experience (Bodovitz, 2008). Interestingly, he claims that the cognitive steps are many separate packages of information which may fade in time in the absence of strong interrelationships; conscious awareness may function as a strong glue that avoids the disintegration of the process of cognition into discrete and independent cycles.

In line with this hypothesis, agents must be aware of what they are doing in order to assign to themselves SoA or SoO or SoR; otherwise, the link between cognition and motivational systems could break down. The congruence between the two sets of signals, i.e., the self-produced intentions of an action (agency) and the sensory signals that arise from their execution (ownership) is the compulsory way of building up the content and attributes of the self. Self-recognition appears early in life because the criterion of congruency has been challenged daily since our birth. It is plausible that this mechanism plays a fundamental role in improving individual knowledge and skill with age.

A number of experiments have been conducted to determine the functional anatomy of the SoA and SoO in normal individuals. For instance, the parietal cortex seems to be the neural region linked to self-perception; while sense of ownership in action execution may be located within the inferior parietal lobe and the temporoparietal junction (Farrer et al., 2003; Ruby & Decety, 2001). According to Jeannerod (2003) SoA and the integrated SoO implies an active organism, i.e. an agent with plans, desires, and actions. Particularly intriguing to us is the self-perception of controlling one's own volitional actions. This statement may well be the necessary link in TBM between the idea of possessing FW and the rise of SoA and SoO.

Moreover, following intentionally caused actions, the sense of agency triggers (in the subject) an interesting illusion concerning the timing of events. By means of Libet's paradigm (Libet, 1983, 2004), Haggard and others demonstrated that when an event is causally linked to a subject's intentional action, the perception of the time separating the decisions from the outcomes of an action is reduced (Haggard, Clark, & Kalogeras, 2002). This sort of binding effect between the two events is strongly correlated to the SoA (Haggard, Cartledge, Dafydd, & Oakley, 2004). Thus, the feeling of exercising FW is fundamental to the sense of self. Altered perceptions of this feeling (generated by hypnosis or by some psychopathological conditions, for instance) may exert an anomalous control of "voluntary" acts, so that the agent reports a distorted perception of the binding effect.

4.5. Point 5

Elsewhere (Bignetti, 2001, 2003), the inherent excitability and firing potential of each single neuron (Katz, 1966) is understood as the intrinsic "desire to think," motivating the neuron to contribute to the thinking process. The expression "desire to think" was provocatively coined for those opposed to the reductionist view of the thinking process. The epistemology of Buddhism considers "desire" to be the insatiable tendency of an individual mind to extinguish all painful stimuli of life (RadhaKrishnan, 1991); again, this can be seen by thermodynamics as any other physical–chemical system which needs to spontaneously evolve to dissipate Gibbs' free energy. However, the question remains as to how the brain can manage the activity of so many neurons in order to be able to execute a goal-directed thought. To identify the mind's "driver" or "organiser" we can either go back to the metaphysical idea of mind–body duality, or try to introduce some type of biophysical mechanism by sorting and integrating a bundle of coherent memories accumulated during the course of a life which may give rise to a virtual personal identity. Scientifically speaking, we prefer this second hypothesis, but from the agent's first-person perspective, the question of self-ontogeny is irrelevant. In a healthy and conscious individual, the performance of a voluntary action in response to outer (environmental) or inner (mental) stimuli is motivated by the "desire" to reach a final goal. It has been

demonstrated that goal-directed polarisation is constantly fed by incentives of various types and values, devised by means of a motivational system (Dickinson & Balleine, 1985, 2002). Extensive psychophysical instrumental training experiments, using rats, have been conducted to understand learning processes. It has been demonstrated that different motivational states may be generated depending on the experimental paradigm applied. In particular, experiments considered both the type of incentive learning, conditioned by aversive and appetitive reinforcement, and the experience of hedonic reactions elicited by action outcome (Dickinson & Balleine, 2002).

The first obvious conclusion we can extrapolate from these experiments is that learning improves as training progresses. Less evident is the mechanism underlying this improvement. Once again, the chemotactic behaviour of the oil droplet in a water maze (point 1) can help us to answer this question (Lagzi et al., 2010). The aim here is not to refer to the 'skill' of the droplet as a paradox, but rather to arrive at a general statement concerning the decision-making process. Every decision must involve both the behaviour of the probabilistic brain and the content of individual memory. According to the basic principles of BDT previously described, the final choice (i.e., the choice of the most likely action) greatly depends on the extent of our knowledge of its effects. The more predictable the effect of an action, the easier it is to make a correct decision and to execute a successful action. Thus, the agent will keep moving passively towards the target, sustained by a driving force that will trace a path of least resistance. Like the droplet in a chemotactic maze, the more coherent and congruent that target appears in our mind, the more efficient our thinking process will be (Bignetti, 2001, 2003). If the affinity between the agent and target is already known, then the action will be the most efficient that can be expected, otherwise, the skill must be acquired by trial and error.

Long ago in some behavioural studies, Tolman demonstrated that voluntary action performance is determined by the incentive value of the outcome of the action itself (Tolman, 1949a, 1949b). In his theory, he introduced the concept of "cathexis" which argued that both animals and humans cannot predict the degree of the success of their actions unless they have already acquired a "cathexis" of what could occur in response to their actions; i.e., they cannot fully predict the intrinsic value of their actions unless they have already tried them. Unlike Pavlovian instrumental learning, Tolman's "cathexis" theory establishes that an unconditioned stimulus cannot automatically trigger a successful response. Thus, the representation of a meaningful incentive value is instantiated in the motivational system as a post-adaptive mechanism. The publication of the "cathexis" theory anticipated Dickinson's work and offered him an extraordinary tool for the interpretation of some experiments performed in rats, where the rats failed to drink sweet drinks when feeling thirsty for the first time due to sudden water deprivation (Dickinson, 1997; Dickinson & Dawson, 1988, 1989). Each motivational system may be fuelled by specific incentive value. An ample variety of behavioural studies have taken advantage of the appetitive behaviour of animals and humans. According to Dickinson and Balleine (2002), behaviour can be learned via two main motivational mechanisms: by the successful outcome of a goal-directed instrumental action, or by the classic conditioning stimuli of aversive or appetitive reinforcement according to the composition of the food.

Every time we act, we have the opportunity to test the relative efficacy of our incentives; thus, we may not only deduce something new about the stimuli, but we may also evaluate the adequacy of our motivational system. In other words, the cognitive processes and motivational systems appear to be linked because depending on the outcome of an action, we learn how to finely tune our motivational system for the future (Bignetti, 2001). In this regard, it is an interesting consideration that FW constitutes a real psychological need of the conscious agent, to the extent that the two things are inextricably linked. The paradoxical element of "intentional" action in TBM is that our knowledge is updated by means of past experience, so we may deduce that cognition is a post-adaptive mechanism. Along the coordinates of knowledge improvement, action will favour cognition and *vice versa* (see Fig. 1). This is a type of feedforward process, which represents one of the most striking examples of the Darwinian evolution of knowledge (Bignetti, 2001, 2004). The mechanism by which we select and accumulate knowledge and skill in our life depends on the cooperation between the UM and the CM. Decision-making and action execution are performed by choosing the best response to a stimulus in memory stores on a statistical basis, but once the action has been performed the UM is unable to evaluate the extent of its correctness. Conversely, the CM cannot decide or perform the action, but it can *a posteriori* evaluate, select and memorise the most correct action from its outcome. Thus, on the one hand, an unconditioned stimulus cannot automatically trigger a successful response; and on the other hand, individuals cannot fully predict the degree of success of an action unless they enact a series of trials and then select and memorise the best one (see the quotation to Tolman's "cathexis" above). In conclusion, the human cognitive theory proposed by TBM is based on *a posteriori* discrimination and the selection of action outcomes similar to the Darwinian evolutionary process that occurs in life sciences, and is in contrast to the Lamarckian-type of cognition theory based on the mirroring of other's actions, proposed by Ramachandran (2010) (see discussion in *point 2*).

5. General remarks

5.1. TBM and ethics

Performing an action the agent is so focused on his OWN first-person perspective that, in that instant, he is

genuinely pervaded by a conviction that he freely decided the right action. If a bit later he were to be assailed by doubts about having made the wrong decision, this thought is already too late, i.e. doubting his own decision is already another story, thus belonging to another action. At best, the agent may rebuke himself for having missed an opportunity. We disagree with Libet (2004) who claims that since the subject's decision is taken too early to be a conscious thought, there is still the opportunity to put a conscious veto; first, because the probabilistic mind promoting the action is unconscious and cannot disagree with itself unless we consider the disagreement still part of the same "decisional" process. Second, the veto (actually, a disapproval) could be conceived as a secondary action only after the subject has observed and evaluated the first action's outcome. A good illustration of this is chensorship during reality TV shows in the US. The increasing demand for live television posed a problem for TV networks because of the potential for technical hitches and inappropriate behaviour and language. The Federal Communications Commission, an independent agency of the United States government, introduced censorship by slightly delaying the broadcast of live programs; this few seconds' delay is sufficient to suppress certain words and images, while keeping the broadcast as "live" as possible. In other words, we cannot put a veto in real time.

The question is, if our actions are decided and executed by the UM who then is legally liable? Let us see, then, how TBM relates to Neuroethics. Neuroethics is a term which was coined in 2002 in the era of applied neurosciences; this discipline combined bioethics and the study of the effect of neurosciences on ethics (Roskies, 2002). In this context, Gazzaniga argues that "personal responsibility is real" (Gazzaniga, 2011) because it is the product of social rules established by people and "is not to be found in the brain, any more than traffic can be understood by knowing about everything inside a car." The accountability of ethical behaviour stands on binomials, such as cause and effect, action and consequence, etc., which belong to a universal architectural principle similar to other information-processing systems (for example, the Internet). Moral rules enable social relationships to be organised on the basis of stable, predictable behaviour in any context and time. Accountability of moral rules in social life provides the automatic brain with a self-protecting servo-mechanism, which may put a veto on decisions that may otherwise conflict with social rules. Although FW is an illusion, we are still responsible for our actions, and brain determinism has no relevance to personal responsibility in real life. To add weight to his arguments, Gazzaniga claims (in a review) that scientific advances in the study of brain mechanisms do not undermine the foundations of the action decision mechanism underlying moral responsibility; so it is time to get over the idea of FW and move on (Gazzaniga, 2012). From a different perspective, Dennet claims that the conclusion that FW does not exist, might means "bad news" (Dennett, 2011). The public generally considers philosophy to be fairly ineffectual in everyday life, however, FW issue matters to people, especially if we consider its role in determining moral behaviour, then philosophers should intervene clearly and unambiguously on the FW issue. Since people may think that FW is a myth, the idea that "my mind made do it" could be a convenient way of passing the buck and escaping blame and penalty. The law presumes 'moral competence' of an individual in order to judge him, then, the main question is whether a robotic mind may acquire a sense of agency and responsibility in order to understand and accept reward or blame. The wiring of our brain circuits provides us with the cognitive ability to bring about the necessary moral competences. Thus, the moral imperative of scientific progress is to discriminate clearly between the circumstances in which an individual can and cannot be considered properly responsible for his action.

Today, neuroethical studies tend to disregard the FW issue, so that whether science demonstrates FW is an illusion or not is irrelevant. This consideration, however, opens up another aspect of mind/body duality. According to TBM, the conscious agent thinks he possesses FW, and thIs belief, though illusionary, is a real and unavoidable part of the individual, thus, the importance of TBM lies in the fact that the first- and third-person perspectives of the role of the conscious agent in intentional action have the same dignity; they serve as tools to understand the mechanism of human cognition. In this mechanism, we do not lose sight of the fundamental role of FW illusion. In this perspective, the fundamental question is: "Is the CM a sheaf of experiences collected and organised by some type of automatism in the brain, or is it the manifestation of a spirit?" If duality does exist it is easier to discuss moral responsibility; however, there is an inherent contradiction in the belief in the automaticity of the brain in intentional actions (FW illusion) and the self-attribution of free responsibility in ethical decisions. Alternatively, we wonder if we can trust the intentions that determine personal and social behaviour if we believe in TBM (see point 3). Conscious FW is invoked to attribute to an individual the responsibility of an intentional action. A man can be liable by law only if his actions have been performed with conscious intentions (*mens rea*) (Morawetz, 1980). According to TBM, FW is not real (at least from the third-person perspective) and thus, the obvious inference is that without FW, we would not have a sense of morality. However, as we perceive SoA and SoR as real, this feeling makes us responsible for determining our moral rules and our compliance with the law (Kahn, 1992). We know from psychology and cognitive neurosciences that moral judgment and intentional behaviour is the result of emotions, affects and rational reasoning ability (Greene & Haidt, 2002). TBM suggests that decision-making and behaviour are the predictable responses to a stimulus chosen from a collection of individual memories sorted by the unconscious mind. The model explains how people falsely believe that they grow up freely and autonomously albeit with cultural

restrictions imposed by the society and the affective and empathic relationships that develop between them and their environment. Since FW illusion is a sort of unconscious error, one is unable to enter into a 'scientific' discussion about it. This belief in FW exists prior to other cognitive process that attempts to disprove it, and thus, TBM will be unable to change the opinion of any individual. However, because laws are acceptable only if their 'meaning' is understood, we can argue that 'education and scholarship' will remain the root of civilisation.

Analysing our theory, we can see that action outcomes and incentives, such as blame and reward, are essential for the conscious mind to learn correct actions. For actions with ethical implications we may consider the motivational incentives of guilt. Feeling guilty may or may not determine an affective state by which one learns how an ethical action should be performed in the future. Moral rules, which are essential for our collective survival, are therefore the product of natural selection. Through socialisation children learn the rules and standards of behaviour are impressed on their memory. This collection of memories could function as a reference library to be utilised by the individual unconscious mind for future actions (point 1 in TBM). Obsessive–compulsive disorder, perpetuated by guilt symptoms that are not easily dispelled, was described by Freud (1929) as the result of a complex struggle of "Ego" against threats from the external world (nature and society), the instinctive demands of "Id" and the critical and moralizing demands of "Super-ego". A malignant super-ego might also be the result of too lenient parenting. Thus, formal education together with familiar and social environments are essential for the imprinting of these moral values.

5.2. Models in the West and in the East: a comparison

We introduced this paper with a quote for Wegner's model on the arousal of FW illusion and apparent mental causation of voluntary action (Wegner, 2002). This model (WWM) originated from an earlier work with Weathley (Wegner & Wheatley, 1999). The main differences between WWM and TBM are in the fundamentals by which a voluntary action is described or in the specific timing of events. First of all, in the WWM, the sequence of events leading to a voluntary action is described as two unconscious processes which bring about conscious thought concerning the action (e.g., intention, expectation) and the voluntary action, respectively. According to the WWM, action 'execution' is delayed with respect to the thoughts that cause it. Thus, causal thought is a sort of explicit prediction of action, which can be validated after execution with the perception of the apparent causal path (which the authors imply most likely gives rise to the experience of will). In TBM, unconscious and conscious mental processes are brain activities with completely different aims that start and intervene at different times, in quick succession. In particular, UM can only elaborate a response to a stimulus thus leading to an action, while CM is activated with the aim of learning and memorising new experiences offered by the relationship between the responses to stimuli and the action outcomes. As already said UM and CM are both brain activity; however CM lags behind UM and has no traces about the UM's activity. The agent's CM erroneously feels as if it is a body-independent entity or soul (primary illusion) who, possessing FW, decides and chooses a voluntary action "free from causes" (secondary illusion). Nevertheless, both illusions turn out to be an inseparable binomial apt for fostering cognition. The originality of this model lies in the causal role of FW illusion, not in predicting or driving the action but in fostering cognition.

Moreover, WWM claims that unconscious mental processes give rise to conscious thought about the action (e.g., intention, expectation). In our opinion, the psychological need in WWM for a conscious mind to decide on the basis of intentions and expectations is a sort of re-emergence of duality, which is often latent in cognitive sciences. The only way to resolve Searle's issue (see above) is to attribute both decision- and action-making completely and exclusively to UM. In TBM, we assume that UM handles both rational and emotional information by means of the same probabilistic mechanism which typically characterises brain activity (Bignetti, 2003, 2010, 2013; Deco et al., 2009; Koch, 1999). On the other hand, one could ask how CM can be motivated by reward/blame incentives in our model. In point 2, we implicitly assume that CM awakening, is accompanied by the experience of a meta-representation of 'ego', the sense of 'self' or 'I'. We will not enter into a discussion of the psychology of the ego, Id and super-ego here, but will assume as true the activation of memory and affective circuits where the neural correlates of motivational incentives such as reward or blame can be found.

Finally, the WWM goes no further than an apparent causal path, which causes the experience of will without explaining whether belief in FW, which is deeply rooted in the psyche, could play a role in conscious processes such as learning and memory. In TBM, FW illusion serves a functional role in connection with both the attribution of responsibility and cognition. We suggest that belief in FW is an unavoidable psychological need to self-attribute a degree of supremacy over nature and that it simply occurs in concomitance with intentional action performance, i.e. an emotional urge for potency. The feeling may wane if the individual is no longer pressured by the urgency of the action and has time to intellectualize it in a detached mood.

TBM has much in common with the epistemology of mind acknowledged by most of the *darshana* of Hindu origin (*Yoga, Advaita Vedanta, Shamkya* and early *Buddhism*), Chinese *Taoism* and Japanese *Zen*. In *Shamkya*, for example, the role of UM is played by '*Prakriti'* (a sort of *natura naturans*) and the role of CM by '*Purusha'* (a sort of thinking self). *Purusha* awakens and is lured by the action of Prakriti and falsely believes he has voluntary decided it (Aurobindo, 2001). As far as Buddhism is concerned, of

particular interest are the teachings of Nagarjuna, the monk of the Mahayana tradition credited with founding the Madyamaka school (approximately 150–250 AD), which claims that sentient beings believe their lives are controlled by volitional actions of a body-independent self, though they are self-less. This is the mistake of the mind leading human beings to duality tied and condemned to a chain of causes and effects which determine the never-ending, painful state of rebirth (*samsara*). Human beings should meditate on the psychological prison created by their own mind to interrupt this endless chain of events and see *Atman* beyond the individual self.

The fact is that in the West we are still debating the nature of self: "Is self a sheaf of experiences collected and well organised by some type of automatism of the brain, or the manifestation of a spirit?" We believe TBM might provide a significant contribution to this debate. However, the correctness of the paradigm as shown in Fig. 1 needs to be investigated further and, to this aim, experiments are currently in progress.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.). *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic press.

Aurobindo, S. (2001). Assays on the Gita. In *The complete work of Sri Aurobindo*. Pondicherry, India: Sri Aurobindo Ashram Editions.

Baars, B. J. (1998). Metaphors of consciousness and attention in the brain. *Trends in Neurosciences, 21*, 58–62.

Bignetti, E. (1994). Vie sensoriali e soft-brain. In *Annals of veterinary medicine faculty* (Vol. XIV, pp. 65–95), Parma.

Bignetti, E. (2001). *Dissacrazione della coscienza. Il Valico Editions*. Firenze.

Bignetti, E. (2003). Cervello e mente: Ovvero casualità e determinismo. In *Annals of veterinary medicine faculty* (Vol. XXIII, pp. 69–78), Parma.

Bignetti, E. (2004). Consciousness can learn but cannot decide. In *Annals of veterinary medicine faculty* (Vol. XXIV, pp. 31–52), Parma.

Bignetti, E. (2010). Free will is the illusionary by-product of self-perception. In *4th International nonlinear science conference*. Palermo: The Soc. for Chaos Theory in Psychology and Life Science (15–17 March).

Bignetti, E. (2013). Ego and free will: A virtual binomial apt for cognition. In *Proc. neural plasticity and cognitive modifiability*, Jerusalem, Medimond Intl. Proc., Monduzzi Ed., 19-22.

Bodovitz, S. (2008). The neural correlates of consciousness. *Journal of Theoretical Biology, 254*, 594–598.

Bullock, T. H. (1970). The reliability of neurons. *Journal of General Physiology, 55*, 565–584.

Daprati, E., Sirigu, A., Pradat-Diehl, P., Franck, N., & Jeannerod, M. (2000). Recognition of self-produced movement in a case of severe neglect. *Neurocase, 6*, 477–486.

Deco, G., Rolls, E. T., & Romo, R. (2009). Stochastic dynamics as a principle of brain function. *Progress Neurobiology, 88*, 1–16.

Dennett, D. C. (2003). *Freedom evolves*. Penguin Group, USA: Viking Press.

Dennett, D. C. (1992b). The self as the center of narrative gravity. In F. Kessel, P. M. Cole, & D. L. Johnson (Eds.), *Self and consciousness: Multiple perspectives* (pp. 103–114). LEA Inc.

Dennett, D. C. (1992a). *Consciousness explained*. Back Bays Books Edition.

Dennett, D. C. (2001). Explaining the magic of consciousness. In *Proc intl symposium on exploring consciousness*, Milan. Fondazione: Carlo Erba.

Dennett, D. C. (2011). "*My brain made me do it*" *(when neurioscientists think can do philosophy)*. Max Weber Lectures (N° 2011/01). Florence.

Dickinson, A., & Balleine, B. (1985). Motivational control of instrumental action. *Current Directions in Psychological Science, 4*, 162–167.

Dickinson, A., & Balleine, B. (2002). The role of learning in the operation of motivational systems. In H. Pashler & R. Gallistel (Eds.), *Stevens' handbook of experimental psychology* (3rd ed. *Learning, motivation, and emotion* (Vol. 3, pp. 497–533). New York: John Wiley & Sons.

Dickinson, A. (1997). Bolles' psychological syllogism. In M. E. Bouton & M. S. Franselow (Eds.), *Learning, motivation and cognition* (pp. 345–367). American Psychological Association.

Dickinson, A., & Dawson, G. R. (1988). Motivational control of instrumental performance: The role of prior experience of the reinforcer. *Quarterly Journal of Experimental Psychology, 40B*, 113–134.

Dickinson, A., & Dawson, G. R. (1989). Incentive learning and the motivational control of instrumental performance. *Quarterly Journal of Experimental Psychology, 41B*, 99–112.

Dietrich, A. (2003). Functional neuroanatomy of altered states of consciousness. The transient hypofrontality hypothesis. *Consciousness and Cognition, 12*, 231–256.

Dietrich, A. (2004). Neurocognitive mechanisms underlying the experience of flow. *Consciousness and Cognition, 13*, 746–761.

Edelman, G. M., & Tononi, G. (2000). *A universe of consciousness. How matter becomes immagination*. New York: Basic Books.

Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., & Jeannerod, M. (2003). Modulating the experience of agency: A positron emission tomography study. *Neuroimage, 18*, 324–333.

Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs an other person as being the cause of an action: The neural correlates of the experience of agency. *Neuroimage, 15*, 596–603.

Freud, S. (1929). *Civilization and its discontent*. Chrysoma Associated Limited (e_books).

Gallagher, S. (2000). Philosophical conceptions of the self: Implication for cognitive science. *Trends in Cognitive Sciences, 4*, 14–21.

Gallese, V. (2007). Before and below "theory of mind': Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of The royal Society B, 362*, 659–669.

Gazzaniga, M. (2011). Who is in charge? *BioScience, 61*, 937–938.

Gazzaniga, M. (2012). *Free will is an illusion but you are still responsible for your actions. The chronicle reviews (The Chronicle of Higher Education)* (July 2013).

Gillett, G. R., & McMillan, J. (2001). Consciousness and intentionality. *Advances in consciousness research* (Vol. 27). Amsterdam, Philadelphia: John Benjamins Publishing Co.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work. *Trends in Cognitive Sciences, 6*, 517–523.

Haggard, P., Cartledge, P., Dafydd, M., & Oakley, D. A. (2004). Anomalous control: When "free-will" is not conscious. *Conscious Cognition, 13*, 646–654.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary actions and conscious awareness. *Nature Neuroscience, 5*, 382–385.

Hallett, M. (2007). Volitional control of movement: The physiology of free will. *Clinical Neurophysiology, 118*, 1179–1192.

Hofstadter, D. R., & Dennett, D. C. (1981). *The mind's I*. Basic Books Inc.

Hume, D. (1739). *A treatise of human nature*. Reprinted By Penguin Classics (1985).

Hurley, S. (2008). The shared circuit model (scm): How control, mirroring, and simulation can enable imitation, deliberation and minreading. *Behavioural Brain Research, 31*, 1–22.

Jacob, P. (2010). Intentionality. In Zalta, Edward N. (Ed.), *The Stanford encyclopedia of philosophy*. <http://www.plato.stanford.edu/archives/fall2010/entries/intentionality/>.

James, W. (1980). *The principle of psychology*. New York: Henry Holt.

Jeannerod, M. (2003). The mechanism of self-recognition in humans. *Behavioural Brain Research, 142*, 1–15.

Jonas, H. (1958). The gnostic religion. *Beacon Press, 1963*, 42.

Kahn, P. H. (1992). Obligatory and discretionary moral judgments. *Child Development, 63*, 416–430.
Katz, B. (1966). *Nerve, muscles and synapses*. New York: McGraw and Hill Book Company.
Koch, C. (1999). *Biophysics of computation*. Oxford, New York: Oxford University Press.
Kording, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences, 10*(7), 319–326.
Krader, L. (2010). In: Levitt, C. (Ed.), *Noetics: The science of thinking and knowing*. New York: Peter Lang Pub Inc.
Kurtz, P. (2003). In B. Karr, & R. Sandhu (Eds.), *Science and religions: Are they compatible?*. Prometheus Book.
Lagzi, I., Soh, S., Wesson, P. J., Browne, K. P., & Grzybowski, B. A. (2010). Maze solving by chemotactic droplets. *Journal of American Chemical Society, 132*, 1198–1199.
Landau, R. S. (2004). *Whatever happened to good and evil?* Oxford: University Press.
Landau, R. S. (2010). *The fundamentals of ethics*. Oxford: University Press.
Lapique, L. (1952). Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation. *Journal de Physiologie et de Pathologie Generale, 9*, 620–635.
Lestienne, R. (2001). Spike timing, synchronization and information processing on sensory side of the central nervous system. *Progress in Neurbiology, 65*, 545–591.
Levinas, E. (1991). *Totality and infinity: An essay on exteriority*. Dordrecht: Kluwer Academic Publishers.
Libet, B. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness potential): the unconscious initiation of a freely voluntary act. *Brain, 106*, 623–642.
Libet, B. (2004). Mind time: The temporal factor in consciousness. In *Perspectives in cognitive neuroscience*. Harvard University Press.
Lingnau, A., Gesierich, B., & Caramazza, A. (2009). Asymmetric fmri adaptation reveals no evidence for mirror neurons in humans. *Proceedings of National Academy of Sciences, 106*, 9925–9930.
Matsuhashi, M., & Hallett, M. (2008). The timing of the conscious intention to move. *European Journal of Neuroscience, 28*, 2344–2351.
Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht: Reidel Pub. Co.
McFadden, J. (2002a). The conscious electromagnetic information (Cemi) field theory: The hard problem made easy? *Journal of Consciousness Studies, 9*, 45–60.
McFadden, J. (2002b). Synchronous siring and its influence on the brain's electromagnetic field: Evidence for an electromagnetic field theory of consciousness". *Journal of Consciousness Studies, 9*, 23–50.
McFadden, J. (2006). The CEMI field theory: Seven clues to the nature of consciousness. In A. Jack (Ed.), *The emerging physics of consciousness* (pp. 385–404). Berlin: Springer (Tuszynski).
Mcguire, P. K., Silbersweig, D. A., & Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain, 119*, 907–917.
Morawetz, T. (1980). *The philosophy of law*. New York: MacMillan Pub.
Norris, D. (2006). The bayesisan reader: Explaining word recognition as an optimal bayesian decision process. *Psychological Review, 113*(2), 327–357.
Northoff, G. (2012). Immanuel kant's mind and the brain's resting state. *Trends in Cognitive Sciences, 16*(7), 356–359.
O'Connor, T. (2013). Free will. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <http://Plato.Stanford.Edu/Archives/Spr2013/Entries/Freewill/>.
Onishi, K. H., & Baillargeon, R. (2005). Do 15 months-old understand false beliefs? *Science, 308*, 255–258.
Radhakrishnan, S. (1991). *Indian philosophy (1 and 2)*. Oxford: Oxford University Press.
Ramachandran, V. (2010). *The neurons that shaped civilization, TED Talks*. <http://www.Ted.Com/Talks/Vs_Ramachandran_The_Neurons_That_Shaped_Civilization.html>.
Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169–192.
Roskies, A. L. (2002). Neuroethics for the new millenium. *Neuron, 35*, 21–23.
Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A pet investigation of agency. *Nature Neuroscience, 4*, 546–550.
Ruch, T.C. (1951). In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley.
Searle, J. R. (1997). *The mistery of consciousness*. The New York Review of Books.
Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience, 11*, 543–545.
Spence, S. A., Brooks, D. J., Hirsch, S. R., Liddle, P. F., Meehan, J., & Grasby, P. M. (1997). A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusion of alien control). *Brain, 120*, 1997–2011.
Steinfels, P. (1995). *Levinas 90, French, ethical philosopher*. <http://www.Nytimes.Com/1995/12/27/World/Emmanuel-Levinas-90-French-Ethical-Philosopher.Html>.
Tolman, E. C. (1949a). There is more than one kind of learning. *Psychological Review, 56*, 144–155.
Tolman, E. C. (1949b). The nature and function of wants. *Psychological Review, 56*, 357–369.
Van Den Berg, R., Vogel, M., Josic, K., & Ma, W. J. (2011). Optimal inference of sameness. *Proceedings of National Academy of Sciences, 109*, 3178–3183.
Van Gulick, R. (2011). Consciousness. In: Edward N. Zalta (Ed.), *The Stanford encyclopedia of philosophy,* summer ed. <http://www.plato.stanford.edu/archives/sum2011/entries/consciousness/>.
Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodiment of mind: Cognitive science and human experience*. Cambridge: MIT Press.
Vohs, K. D., & Schooler, J. W. (2008). The value of believing in free will. *Psychological Science, 19*, 49–54.
Von Hofsten, C. (2004). An action perspective on motion development. *Trends in Cognitive Sciences, 8*(6), 266–272.
Wegner, D. M. (2002). *The illusion of conscious will*. Massachussetts Institute of Technology.
Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist, 54*, 480–492.