Behavioral Sciences and the Law
Behav. Sci. Law **25**: 221–234 (2007)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/bsl.754

Free Will, the Self, and the Brain

Gilberto Gomes, Ph.D.*

The free will problem is defined and three solutions are discussed: no-freedom theory, libertarianism, and compatibilism. Strict determinism is often assumed in arguing for libertarianism or no-freedom theory. It assumes that the history of the universe is fixed, but modern physics admits a certain degree of randomness in the determination of events. However, this is not enough for a compatibilist position—which is favored here—since freedom is not randomness. It is the *I* that chooses what to do. It is argued that the core of the free will problem is what this *I* is. A materialist view is favored: The *I* is an activity of the brain. In addition to absence of external and internal compulsion, freedom involves absence of causal sufficiency of influences acting *on* the *I*. A more elaborate compatibilist view is proposed, according to which causal determination is complete when we add events occurring *in* the *I* (of which the subject is not conscious). Contrary to what several authors have argued, the onset of the readiness potential before the decision to act is no problem here. The experience of agency is incomplete and fallible, rather than illusory. Some consequences of different views about freedom for the ascription of responsibility are discussed.

The problem of free will has been an object of human concern since antiquity. In the first section of this article, the problem is outlined and three possible solutions (no-freedom theory, libertarianism and compatibilism) are examined. The next section shows the importance of the concept of determinism in the evaluation of these alternative views. A defense of compatibilism will then be presented. It will be shown that the *I* is a key concept in evaluating the plausibility of compatibilism. However, the usual view of compatibilism is vulnerable to criticisms and a more elaborate view of this position will be proposed in the fourth section. The causal status of the *I* in relation to the possibility of doing otherwise is discussed next. The sixth section will briefly focus on freedom as an inevitable consequence of the normal

*Correspondence to: Gilberto Gomes, Ph.D., R. Lopes Quintas 100-605-I, 22460-010 Rio de Janeiro, RJ, Brazil. E-mail: ggomes@uenf.br

working of the human brain. Some data from neuroscience has been invoked as evidence that it is not free will that initiates voluntary acts. This claim, which has been used to favor either no-freedom theory or special forms of libertarianism, is countered in the seventh section. Psychological experiments have also been used to try to prove that free will is an illusion, and these are succinctly discussed in the eighth section. The last section examines some consequences of the different positions concerning free will for our views about responsibility.

THE PROBLEM AND THREE POSSIBLE SOLUTIONS

Are human beings free? Should human beings be free? A little reflection will show that the word *free* is used in two different senses in these two questions. The question of whether human beings *should* be free implies that they may *not* be free under certain circumstances. It involves a sense of freedom that refers to external circumstances. A dictatorship, for example, severely restricts the freedom of its citizens, who are thus not free to do several things that they should be able to do. The question of whether human beings *are* free is a question of whether they have a certain intrinsic property, namely, the capacity to choose among the alternatives that are available to them, in a way that is influenced, but not completely determined, by all sorts of factors that affect them. We might call these two senses *external* and *internal* freedom, respectively.

The problem of free will may be conceived as the problem of whether we have internal freedom, what this internal freedom consists in, and how it relates to the notion that all events have causes. Since it is usually believed that only agents that have internal freedom may be considered responsible for what they do, these questions may have significant implications for social life, moral judgments and the legal system.

The following are three possible views concerning the problem of free will, as it has been discussed in recent times.

1. There is no free will. Free will is an illusion. What a person does is determined by what goes on in her or his brain, which is in turn determined by previous internal and external events. The causal chain goes back to times when the person did not yet exist, so the person cannot be held ultimately responsible for what she or he does. Therefore, legal punishment can only have a practical justification, not a moral one. If someone is good or bad, this is just the result of the circumstances of her or his environment and of her or his nature. Even if the person did something good as a result of an effort to do so, this is because she or he was lucky enough to realize that the effort was worth making and was lucky enough to have a nature that enabled her or him to make it. When someone wonders or ponders what she or he will or should do, the result of this decision-making process is already determined by her or his nature and the present circumstances. The person simply does not know what this result will be, but she or he is not in fact free to determine it. It is an illusion to think that what the person will do depends on her or his decision, because, even if it does depend on a particular decision, the decision itself is determined by prior events and not by a prior decision. There is nothing a person can do to change the course of events, including her or his own

DOI: 10.1002/bsl

actions. Concerning past actions, it is false that the person could have done otherwise, if she or he had decided to do so. All a person did is what she or he had to do, since it was determined by her or his nature and her or his circumstances at the time of the action, which were not ultimately determined by her- or himself.

2. It is an evident fact that people are free to choose what they do. There are practical limits to human action, but whenever more than one course of action is possible, people can choose which one they will take. Physical determinism cannot be extended to human actions. The will causes physical events but it is not caused by physical events. When choosing what to do, people are influenced by all sorts of internal and external factors, but these factors do not completely cause the decision. Even if a person's decision is influenced (that is, partly determined) by factors of which she or he is not conscious, there are other factors of which she or he is conscious and, in relation to these, it is the person who chooses, at least to a certain point, how and to what extent they will influence her or his decision. Therefore, free will cannot be reduced to neural processes in the brain. These are physical processes subject to causal laws. If the will were a part of these, it would be determined, not free. Free will is something that can change the natural course of events.
3. We know that some actions are freely chosen and others are automatic reactions. We should be able to able to characterize the difference between them. Science has shown us that all natural phenomena are caused by other natural phenomena. The activity of the human mind has been consistently shown to depend on the functioning of the human brain and the latter consists of natural phenomena involving neurons, action potentials, neurotransmitters and so on. There is no reason to suppose that the causal chains that determine neural events are broken at some point. However, we should not deny that free will exists, since we need this concept to distinguish between two sorts of human action. Therefore, we must understand free will as a part of the workings of the human brain. We should change our concept of freedom, so as to preserve what is essential in it, eliminating the idea that freedom must escape natural causality. An action is free when it results from a conscious intention to start it, to continue it, or at least not to refrain from starting or continuing it. There is no need to suppose that this intention must not be the result of causal processes in the brain. A person is responsible for an action when this action results from a conscious intention. To be responsible for an action is not to be *ultimately* responsible for it, in the sense of also being responsible for all the events in the causal chains that led to the existence of the conscious intention that determined the action.

The three statements above illustrate the three main alternative positions on the problem of free will as it is nowadays discussed. The first two of these admit that there is an incompatibility between freedom and natural causality. Thus, they are both *incompatibilist*, but they derive opposite conclusions from this incompatibility. The first maintains that all events are causally determined and consequently there is no free will. Following Galen Strawson (2004), I will call this position the '*no-freedom theory*'. The second accepts that free will exists and consequently rejects that all events are causally determined. This position is usually called *libertarianism*. Some versions of it admit that events may be caused either by other events (natural causality) or by agents (agent causality). However, the causing of an event by an

DOI: 10.1002/bsl

agent is not itself considered an event that is caused by other events. The third position admits that freedom and natural causality are compatible. It is thus called *compatibilism*.

There are many versions of these three basic positions. Libertarianism may be associated with a dualist conception, according to which physical reality and the mind (or soul) are essentially different realities (or substances) that interact with each other. However, there are also non-dualist versions of it, though these have difficulty in showing that they do not really imply some form of dualism. Different authors may draw different moral, social and legal implications from no-freedom theory, and at least since Hobbes several different compatibilist definitions of free actions and free will have been proposed.

The aim of this article is not to examine the different versions of the three positions and the arguments for and against each of them. Rather, it will argue for a particular version of compatibilism, which tries to reconcile the no-freedom theorist's view that human action is naturally caused with the libertarian's claim that there are actions that are chosen by the person rather than determined by causal factors acting on her or him.

THE CONCEPT OF DETERMINISM IN DISCUSSIONS OF THE FREE WILL PROBLEM

In discussions of the problem of the relation between free will and universal causality—i.e. the idea that everything that happens (including actions) is caused by prior events—causality is often understood as strict determinism. Strict determinism is the idea that prior conditions precisely determine every parameter—in minute detail—of what happens. Nothing is indeterminate; there is no essential randomness in nature. Randomness, according to this view, is just ignorance of causes. When we toss a coin or throw a dice, a precise knowledge of the relevant parameters, together with the laws of mechanics, would allow a precise prediction of the result. This conception of the world was specially stimulated by classical mechanics. Here we have the so-called universe of Laplace, who claimed that an intellect that knew all the forces acting in the universe and the position of all its objects at a given instant would be able to know all its past and future states (Laplace, 1825/1921).

The advent of quantum mechanics, however, has significantly changed this picture of the world. In quantum physics, some events have only a definite probability of occurrence—they are not strictly determined. Strict deterministic laws apply only at the macroscopic level, not at the microscopic one. However, microscopic events may have macroscopic effects, as chaos theory and the study of self-organizing systems have shown (Prigogine, 2003). Therefore, events that involve the evolution of dynamic systems far from equilibrium are not fully predictable from prior conditions.

Even at the intuitive level, strict determinism may seem to be a strange doctrine. On the one hand, the idea that everything is strictly determined is attractive. On the other hand, however, consider a grain of dust that falls somewhere on a desert planet at a certain moment: It seems bizarre to think that since the Big Bang it was determined that precisely this grain of dust would fall exactly on this place at this very moment! In other words, it is not only in relation to free actions that strict

DOI: 10.1002/bsl

determinism seems unlikely, but also in relation to many natural events that involve no agent at all. Moreover, since modern science does not say that strict determinism exists in nature, why should we bother with the intuitive incompatibility between free will and strict determinism?

The argument above is given here because many discussions of free will assume strict determinism as an established property of physical events. They often imply that any non-libertarian theory has to admit the strict determinism of human actions, which is not true. Galen Strawson (2004), for example, states

> According to *compatibilists*, we do have free will. They propound a sense of the word 'free' according to which free will is compatible with *determinism*, even though determinism is the view that the history of the universe is fixed in such a way that nothing can happen otherwise than it does because everything that happens is necessitated by what has already gone before (p. 2).

This makes it seem as though the only option for compatibilists is to adopt strict determinism. This is not so: Compatibilists may endorse a view of causality that allows for a certain degree of randomness in the determination of events. However, the view that nature is not strictly deterministic is not enough to offer a solution to the problem of free will. To admit a certain degree of randomness in the determination of an action is insufficient to characterize it as free and to admit that one's present nature is the result of a process of development in which many events involved a certain degree of randomness is also of no avail to make one responsible for one's actions. A free action is felt to have been determined by the person who does it, not by chance. A random event is not one for which a person would be held responsible. Therefore, one might argue that if a person cannot be held responsible either for chance or for prior circumstances, she or he cannot be held responsible for her or his actions, if prior circumstances and chance are all there is to the determination of actions.

A DEFENSE OF COMPATIBILISM—HIGHLIGHTING THE ROLE OF THE *I*

Libertarians feel that the person her- or himself is the essential element in the determination of free actions—and I believe they are right in this. This leads us to ask: What is the person her- or himself? What is the *I* (or self) that decides an action, when a person can say, "I did it because *I* decided to do it"? The concept of *I* is an elusive one and to say it is an indexical[1] is not enough to dispel this elusiveness. Suppose I cut my finger. I can say, "*I* have been hurt", but I can also say "*My finger* has been hurt". It would be easy to observe that my finger is just a part of myself, but I can also say, "*I* feel an ache *in my finger*", and in this case the *I* seems not to include the finger—the finger is in some sense external to the *I*, since the *I* is what feels and the finger is what is felt. The finger is a (bodily) part of the person that is felt by another (mental) part of the person. Note also that we say "*my* finger" just as we say "*my* pen", as if they were both things that belong to the *I*, not only the pen but also the finger is in some sense external to the *I* that possesses them.

[1]An indexical is a term whose reference depends on the context of utterance. *I* is an indexical because it refers to whoever is speaking.

DOI: 10.1002/bsl

One might say that the innermost *I* is the psychological part of the person, as distinct from the body that she or he possesses, feels, and moves and in which her or his emotions express themselves. However, one can say, ''*I* had *my reasons* for doing this'', including feelings, ideas and plans among these reasons. So these reasons are psychological factors that in a certain sense are *in* the *I*, but in another sense act *on* the *I* and are in this sense external to it. The *I* considered these reasons, thought about how they should be weighed, and then decided what to do. One may say that it was not these reasons that determined the action, but the *I*. The *I* that considers the reasons is in a certain sense distinct from them—but what is it?

I think that the core of the problem of free will is the problem of what the *I* that chooses what to do is. From a strict Humean point of view, an *I* that is distinct from reasons, feelings and plans simply does not exist, since Hume conceives the mind as a mere bundle of sensations, along with thoughts and intentions derivs from them. A radical empiricist[2] position recognizes no active center in the mind—the mind is essentially passive, since it depends completely on sensation of what is external to it. Such a viewpoint agrees well with no-freedom theory. In fact Hume defends a sort of compatibilism. On the one hand, he holds that actions and volitions, no less than physical events, obey a strict determinism. On the other, he grants that humans have ''the power of acting or not acting according to the determinations of the will''—which is his definition of liberty. Moreover, he thinks that liberty is compatible with determinism (Hume, 1748/2000, p. 72).

However, Hume's compatibilism is achieved through a very restrictive view of freedom (liberty), according to which liberty is no more than what is ''allowed to belong to everyone who is not a prisoner and in chains'' (Hume, 1748/2000, p. 72). Moreover, the ''will'' mentioned in the quotation above is subsequently analyzed by him as being nothing more than motives, inclinations and circumstances—not an active power that evaluates these motives, inclinations and circumstances and determines what to do. Hobbes (1654/2005, p. 273) also conceived freedom as the mere ''absence of all impediments to action that are not contained in the nature and intrinsical quality of the agent''. This corresponds to what I called external freedom at the beginning of this article.

Criticisms of the compatibilist position often assume that it adopts such a restrictive view of freedom as the absence of external restraint. Searle (2000), for example, states

> The compatibilist view is that if we properly understand these terms, freedom of the will is completely compatible with determinism. . . . So if someone puts a gun to my head and tells me to raise my arm, my action is not free, but if I raise my arm by way of voting, as we say, 'freely', or 'of my own free will', then my action is free. Though in both cases . . . my action is completely causally determined (p. 11).

Searle goes on to argue that this sense of freedom as absence of external constraints is irrelevant to the problem of free will. Moreover, even if we add the absence of *internal* constraints such as compulsion, panic, or addiction, the critic of compatibilism will remain unconvinced. According to Searle, we experience a gap between the reasons for acting and the decision to act. We feel that the antecedent causal conditions of our free actions are not causally sufficient to produce the action.

[2]Radical empiricism is the philosophical position according to which all mental events are completely derived from sense impressions.

DOI: 10.1002/bsl

Accordingly, freedom seems to imply not only absence of external constraint and internal compulsion, but also absence of causal sufficiency of the antecedent conditions of an action.

Consider another quotation from Searle (2000, p. 12):

> A complete specification of all the psychological causes operating on me at t_1, with all their causal powers, including any psychological laws relevant to the case, would not be sufficient to entail that I would perform act *A* under any description.

Note that the author uses the preposition '*on*' when referring to the relationship between the psychological causes and the *I*. The psychological causes he mentions act *on* the *I* and are thus in a certain sense external to the *I*. We see that the author is considering the *I* or 'me' as being distinct from all the psychological causes at play. Suppose now we also take into account the psychological causes and laws operating *in* me (and not only those operating *on* me, as Searle does). Then a complete specification of all the psychological causes and laws acting both *on me* and *in me* might be sufficient to entail at least a certain probability of my performing act *A*.

What I want to suggest is that we should adopt a more complex view of the *I*. There are multiple *I*s, according to different aspects or moments of mental activity. When I say that I have a certain desire, it is certainly I who desire, but when I choose to resist this desire, it is another *I* that is active. The libertarian is right when she or he says that the reasons and influences that operate on the deciding *I* are insufficient to determine the action, but this leaves open the question of whether what the deciding *I* actively does with these influences in order to reach a decision is itself determined by causal factors or not.

When the libertarian says, ''It is I, and not antecedent causes, that determined what I did'', she or he is not considering the possibility that her or his *I* is included in these antecedent causes and is itself caused by previous conditions. That this may be the case should be considered as a scientific hypothesis, not to be simply rejected on philosophical or phenomenological grounds. Scientific facts often go against how things appear to be.

Here the no-freedom theorist might step in and say, ''But if the 'I' is causally determined, then it is not free.'' On behalf of compatibilism, I would like to counter this objection with the argument that it all depends on how we define freedom. If we include in the definition of freedom the requirement that free decisions to act must not be causally determined, then compatibilism would be simply impossible. Perhaps this is the most natural view of free will. Perhaps we tend to view free decisions to act as emanating from the *I* and the *I* as being outside the natural sequences of causes and effects, but the latter view may be wrong even if the former is right.

A MORE ELABORATE COMPATIBILIST VIEW OF FREE WILL

If one's starting point is a conception of free will that includes its not being causally determined by prior events, then any compatibilist proposal will involve the need to change this conception. However, such a reconceptualization of free will need not be so restrictive as to limit freedom to the absence of external constraint (external freedom), as critics of compatibilism usually imply. A more elaborate compatibilist

DOI: 10.1002/bsl

concept of free will may indeed include the libertarian's point that the causal factors acting *on* the deciding subject do not completely determine her or his actions (Hodgson, 2005). I believe that a compatibilist conception may preserve all that is essential in the everyday notion of free will, namely (1) the idea that in free action the person her- or himself, as a psychological subject, chooses what to do, (2) the idea that freedom depends on consciousness and, more specifically, that a free action results from a conscious intention, and (3) the idea that freedom implies the possibility of doing otherwise. Let us examine these three points.

1. The idea that it is the person that chooses what to do is fully consistent with natural causality, if we adopt the ''astonishing hypothesis'', as Crick (1994) calls it, that all our personhood and psychical life is given by the activity of our brains. The working brain, with literally billions of synaptic contacts among its neurons, is such a wonderfully differentiated and complex system that it has been increasingly considered reasonable to admit that all our mental life corresponds to its activity. Accordingly, the brain processes of considering alternative courses of action and choosing what to do are also to be regarded as subject to natural causality.
2. Consciousness (including conscious intentions) may also be considered as a property of complex neural activity (as argued, for example, by Gomes, 1995).
3. The possibility of doing otherwise has often been judged to be one of the hallmarks of free action. One has done something of one's own free will when one could also have done otherwise. Searle (2000), for example, states

> Granted that the action did occur, and that it did occur for a reason, all the same, the agent could have done something else, given the same causal antecedents of the action (p. 11).

We have already seen that in such a line of reasoning the agent her- or himself is not included among the causal antecedents that are being considered. Compatibilism can grant that, given the same causal antecedents *other than the agent's activity of deciding what to do*, she or he could have done something else. Now this activity of deciding what to do need not be viewed as something outside the realm of natural events. It may also be a natural event that is the result of causal conditions. This naturalistic conception does not eliminate the possibility of doing otherwise.

What do people (including philosophers) mean when they say that someone could have done otherwise? They surely do not merely mean that the person concerned might have done something different from what she or he did, if circumstances had been different from what they were. Any event—not just free actions—might have been different from what it was if its causes had been different. What they mean is that the person concerned had the *power* to do otherwise even if circumstances had been the same. If the person had that power, it would have been possible for her or him to do something else in the presence of the same causal factors acting on her or him. This means that what the person did was not completely determined by *these* causal factors. But again, the circumstances imagined to be the same do not include the person's very process of deciding what to do. (The theme of the possibility of doing otherwise is further developed in the next section.)

In a nutshell, according to this more elaborate version of compatibilism, a free action is at the same time (1) free of complete determination by conditions external

DOI: 10.1002/bsl

to the system of the person's mind that makes the conscious decision to act and (2) causally determined by conditions internal to this deciding system.

THE POSSIBILITY OF DOING OTHERWISE

An action is free (or is determined by free will) if the person who did it could have done otherwise, even if all the causal factors acting on her or him had been the same. The emphasis on the role of the *I* in the discussion above should in no way be considered as an alternative to this traditional way of characterizing free will. The presence of an *I* is not sufficient for an action to be free. An action may be due to a pathological compulsion, for example. Such an unfree action may be said to be a manifestation of a person's *I*, since the word *I*, as discussed above, may be used to refer the whole of a person or to different physical or mental subsets of her or him. The point in highlighting the role of the *I* is that the concept of an *I* is implicitly present in the traditional conception of free will as involving the possibility of doing otherwise.

In order to see this, it is necessary to articulate what is implicit in the usual formulation of this conception. People usually simply say that an action is free if the person could have done otherwise. However, this is patently not enough. As pointed out above, it is implicit that this must have been so *even if the circumstances had been the same*, but this is still insufficient. If someone argues that a person could have done otherwise in the same circumstances if she or he had had different desires or different beliefs, the advocate of free will not agree. The idea of free will seems to imply that a person could have done otherwise in the same circumstances *even if she or he had had the same desires and beliefs*. One might say, ''My desires and beliefs have influenced my decision, but they have not *determined* it. *I* considered these desires and beliefs and *I* decided what to do. *I* could have done otherwise, even if I had had the same desires and beliefs.''

We thus find that the *I* is an essential element in the idea of the possibility of doing otherwise. In the reasoning above, as argued in the third section, the desires and beliefs seem to be *inside the person*, but still *outside the deciding system in her or his mind*—referred to as her or his *I*. When one says, ''*I considered* my desires and beliefs . . .'', one is speaking about an active instance in one's mind that exists *in addition to* her or his desires and beliefs and is able to do something with them. The fundamental question then is how to conceive this *I*. We tend to view this *I* as lying outside the realm of physical causes and effects, but is this really so? My compatibilist proposal is that we should view it as a system in the brain. We can then keep the idea that this system could have chosen to act otherwise, even if all the circumstances, beliefs, and desires that it was considering had been the same. However, it is not logically possible that it could have chosen to act otherwise if everything that occurred *in it* had been the same, since the choice of an action is a direct result of what occurs in it.

NOT FREE NOT TO BE FREE

Sartre (1943, pp. 168, 494) provocatively stated that man is condemned to be free. A person is free to choose no longer to live (by committing suicide), but one is not free

DOI: 10.1002/bsl

to choose no longer to be (internally) free. It is in the nature of normal human beings to have the capacity to determine by themselves which available course of action to take and it is not in their power to give up this capacity.

From the point of view of neuroscience, this may be taken as meaning that the human brain is naturally endowed with an ability to consider different possibilities of action and to choose one of them. Moreover, this ability seems to be essential to personal identity, since there is a sense that the same self is the agent of different instances of deciding what to do, in a person's life. Thus, the identity of a person, both to her- or himself and to others, is intimately bound with this unavoidable ability to act freely and the innumerable instances in which it is used.

NEURAL PROCESSES INVOLVED IN VOLITION AND THEIR IMPORT FOR THE FREE WILL PROBLEM

Many have interpreted the results of Libet's experiments on the timing of the conscious decision to act (Libet, Gleason, Wright, & Pearl, 1983a; Libet, Wright, & Gleason, 1983b) as evidence that free will is an illusion (Wegner, 2002), or at least that it is ineffective as regards the *initiation* of free acts (Libet, 1985, 1987). These experiments involve the *readiness potential* discovered by Kornhuber and Deecke (1965). The readiness potential is obtained by averaging a certain number of electroencephalographic (EEG) tracings: A slow negative potential appears mainly over the motor cortex half a second or more before a voluntary movement is performed. The movement is previously defined by the experimenter, but the moment of making it is determined by the subject, who is instructed to perform it immediately after the decision.

The interval between the onset of the readiness potential (RP) and the movement is influenced by the instructions given. It may be longer than a whole second. Libet, Wright and Gleason (1982) were able to obtain shorter RPs with instructions that favored spontaneity. However, the RPs still preceded the movement by at least 200 milliseconds (ms) and typically by about 600 ms.[3] This seems to be long in relation to the experienced time between the decision to move and the movement. The correct conclusion to be drawn from this is that the conscious decision to act occurs after the onset of the RP. However, the relatively long interval between the RP onset and the movement has often led to a less parsimonious conclusion that may be stated in the following terms: ''The brain initiates the neural events that produce the movement before the mind decides to make the same movement. Consequently, it is the brain and not the mind that initiates the movement. Neural events occurring before a conscious decision cannot have been determined by this decision. They are determined by prior neural events. Therefore, they are not determined by free will. The subject feels that the moment of moving was freely chosen by her- or himself, but in fact it was determined by events that preceded the conscious decision.''

This kind of conclusion is coherent if we adopt a dualist view of the relation between mind and brain. The RP poses a serious problem for those who think that free will is an attribute of an immaterial mind that is not subject to natural causality. Eccles, a distinguished neurophysiologist who advocated such a view, devised an

[3]For a thorough discussion of the data, see Gomes, 1998.

DOI: 10.1002/bsl

elaborate neurophysiological hypothesis to account for the temporal relation between the onset of the RP and the conscious decision to move (Eccles, 1985). This hypothesis really offers a solution to the problem, but there is no evidence to support it (Gomes, 2005).

For a monist[4] view of the mind as the complex activity of the brain and a compatibilist view of free will, however, no such problem exists and no conclusion such as that stated above is justified. The conscious decision to make the movement at a certain moment is also a neural event that is caused by prior neural events. There is no sense in saying that it is the brain and not the mind that initiates the movement, since the mind is an activity of the brain (Bunge, 1979). The fact that the decision is free means that it was determined by the person's conscious *I* or self and that this conscious *I* could have chosen another moment to make the movement. However, since the conscious *I* is a neural system in the brain, there is no incompatibility between this decision being free and the fact that it was preceded by certain observable neural events (manifest in the initial part of the RP). These may be seen as a part of the workings of this neural *I* (Gomes, 1999).

There is no reason to suppose that the conscious decision to immediately make the movement arises out of nothing. That the subject is not conscious of its causes is no reason to suppose that it has no causes. Moreover, to suppose that it has causes is no reason to suppose that it is not free, as argued in the defense of compatibilism given above.

This sort of confusion is apparent in the work of authors who are not themselves dualists concerning the mind–brain relation. Haggard (in Haggard & Libet, 2001, p. 50), for example, states that ''[t]emporal precedence may be important for the controversy between mind-to-brain vs brain-to-mind causation''—but this is true only from a dualist perspective. And again: ''the long gap between RP onset and [the awareness of the decision to move ...] justifies a brain–mind rather than a mind–brain direction of causation'' (p. 53). He continues: ''The free will theorist could suggest that conscious intentions cause the brain processes of movement selection ...''. But the *compatibilist* free will theorist could suggest that conscious intentions *are* brain processes that cause these other brain processes of movement selection and are caused by still other brain processes that precede them. To speak of mind-to-brain or brain-to-mind causation implicitly involves a dualist view of the mind–brain relation.

In sum, from a monist and compatibilist perspective, the fact that a conscious decision is caused by (other) neural events is no reason to believe that voluntary acts are not initiated by free will.

PSYCHOLOGICAL EVIDENCE AGAINST FREE WILL?

A different kind of evidence was presented by Wegner (2002) as supporting the view that free will is an illusion. In one of his experiments, subjects see someone else's arms in a mirror, in a position that makes them look like the subject's own arms. After hearing a command to make a certain movement, they see the hands in the mirror

[4]Concerning the mind–brain relation, dualism is the view that mental events and physical events are essentially different realities, while (materialist) monism is the view that mental events are brain events.

DOI: 10.1002/bsl

making that movement. This makes them feel somewhat as though they had made the movement. Wegner argues that thoughts and actions alike are unconsciously caused. When the conscious thought of an action precedes consciousness of such an action, the subject infers that the thought caused the action, but this is an illusion: Both were unconsciously caused (Wegner & Wheatley, 1999).

Nahmias (2005) has adequately discussed the conceptual confusions present in Wegner's theory. The conditions identified by Wegner for producing the experience of agency are not sufficient: Subjects report a slightly enhanced feeling of control but they do not really think they have made the movement. Neither do they seem to be necessary, since no specific thought of the action precedes highly automatic fast actions such as occur in sport or musical performance, but these are still felt by the subject to have been made by her- or himself. Furthermore, the fact that the experience of agency is not infallible, particularly in pathological conditions, does not imply that it is systematically illusory. As Nahmias (2005, p. 777) remarks, "here it is more appropriate to describe our experiences as *incomplete* rather than illusory".

FREE WILL AND RESPONSIBILITY

What are the consequences of different views about free will for the ascription of responsibility or guilt for one's actions and for the justification of punishment? Libertarian incompatibilists tend to exaggerate the autonomy of the subject in the determination of her or his actions. As they think that choices are made by the self without being causally determined, they view the self as the ultimate and absolute originator of free actions. One's actions emanate from one's self, and though the self may change, it alone is the originator of any such change. Even if there was external influence on the determination of an action or of a change in the self's character, it is the self that decided to accept this influence and this decision has not been caused. The idea of self-origination that is present in such a conception is hardly intelligible, as Nietzsche (1886/1973) has eloquently contended. A may be the origin of B, but what is the meaning of saying that A is the origin of A?

Libertarians allow, of course, that the self may not be free in certain cases. Mental illness is one of them. Young age is another. They usually think that what one does in these cases cannot be considered as having been freely chosen by the self, and consequently one is not responsible for one's actions. The problem is that the libertarian conception of an absolute free will favors an all-or-nothing contrast between free and unfree actions. Within this view, it is difficult to conceive of gradations or transitions regarding the ability to choose freely. It seems that a free action's escape from causality is either present or absent. We obtain a black and white description with no shades of gray. However, there is a smooth transition from childhood to adolescence to adulthood. Pathological compulsions also vary in degree and this suggests a correlative variation in the degree of free will.

No-freedom theorists often question the existence of moral responsibility and the justification of punishment on moral and not only practical grounds. They reason that if actions are determined by causes, then the subject is not responsible for them. This leaves us with an incomprehensible picture of the human world, since there is no responsibility or moral obligation in it. If one could not have done otherwise, it cannot be the case that one ought to have done otherwise (Howard-Snyder, 2006).

DOI: 10.1002/bsl

In everyday life, there are obviously cases in which we consider someone responsible for what she or he did, and others in which we consider her or him not responsible for it. It cannot be that the difference between them is merely that we are victims of an illusion regarding the former and we are not victims of such an illusion regarding the latter.

No-freedom theorists often justify punishment on a mere utilitarian basis. Although a person is never responsible for what she or he did, they argue, there are cases in which punishment is useful to avoid future unhappiness, and others in which it is useless. Punishment is justified regarding the former. However, why should the presence or absence of an illusion in the observer make a difference in the usefulness of punishment of the person who made the action? The no-freedom theorist may answer that what really makes a difference is the presence or absence of a certain attribute in the agent, just before the action, and that the presence or absence of this attribute correlates with the presence or absence of the illusory attribution of responsibility by the observer (which may be someone else or the agent her- or himself). When this attribute is present in the agent, punishment shows itself to be useful and observers tend to have the illusion that the agent was responsible for the action. This is to say that, even though the agent is not responsible, there is something about her or him, in certain cases, that makes punishment useful.

The definition of just what this attribute is may be problematic, but the no-freedom theorist will not allow that it may be called responsibility, since she or he has already concluded that responsibility does not exist. I believe it is much more reasonable to *define* responsibility as being precisely this attribute, whatever it may turn out to be and however difficult it may be to ascertain it. If the word *responsibility* is so defined that it cannot be a property of the agent, but at the same time *there is* a property of the agent that is relevant both for what people think and for the usefulness of punishment, then the word should be *redefined* so as to refer precisely to this property.

A change in the definition of *responsibility* used by certain authors is precisely what compatibilism requires and supports. Someone need not be the ultimate originator of an action, in the sense of being an uncaused cause or ''prime mover unmoved'' (Chisholm, 1982), to be responsible for it. An action is free, and the agent responsible for it, when it results from a conscious *decision-making* process and the agent could have done otherwise if she or he had decided to do so. The fact that the person's *decision-making* process must have had causes and that a different decision could only have been made if a different set of such causes were present should not be considered as removing her or his responsibility. A person is responsible for an action when her or his *I* was in control of the process of deciding to make this action.

According to the compatibilist view I am defending, actions are free and the person is responsible for them when they derive from the person's *I* and this *I* could have chosen to do otherwise, but this *I* is not an abstract or supernatural entity outside the realm of natural causality. The *I* is a self-organizing and self-steering system (van Duijn & Bem, 2005) within a brain. It is not a merely passive reflection of external influences. It has individuality (personality) and consistency over time, though it is subject to change. Change in the *I*'s character is usually slow, but in exceptional cases (including those of religious conversion) a large change may occur at a particular moment or over a short period of time. Moreover, decisions are not made by an impartial and purely rational decision-making system, but rather by a

DOI: 10.1002/bsl

motivated and emotional one. A compatibilist theory thus seems better equipped to account for the causal factors that may act *in* the *I* without the person being conscious of them. Such a psychological or neuroscientific explanation, however, does not deprive a person of her or his responsibility for actions that have been chosen by her- or himself from among more than one alternative.

REFERENCES

Bunge, M. (1979). *Treatise on basic philosophy* (Vol. 4) Dordrecht: Reidel.

Chisholm, R. M. (1982). Human freedom and the self. In G. Watson (Ed.), *Free will* (pp. 24–35). Oxford: Oxford University Press.

Crick, F. (1994). *The astonishing hypothesis: The scientific search for the soul*. London: Simon and Schuster.

Eccles, J. C. (1985). Mental summation: The timing of voluntary intentions by cortical activity. *Behavioral and Brain Sciences, 8*(4), 542–543.

Gomes, G. (1995). Self-awareness and the mind–brain problem. *Philosophical Psychology, 8*(2), 155–165.

Gomes, G. (1998). The timing of conscious experience: A critical review and reinterpretation of Libet's research. *Consciousness and Cognition, 7*, 559–595.

Gomes, G. (1999). Volition and the readiness potential. *Journal of Consciousness Studies, 6*(8/9), 59–76.

Gomes, G. (2005). What should we retain from a plain person's concept of free will? *Journal of Consciousness Studies, 12*(1), 40–43.

Haggard, P., & Libet, B. (2001). Conscious intention and brain activity. *Journal of Consciousness Studies, 8*(11), 47–63.

Hobbes, T. (1654/2005). Of liberty and necessity. In *The English Works of Thomas Hobbes of Malmesbury* (Vol. 4). Boston, MA: Adamant.

Hodgson, D. (2005). A plain person's free will. *Journal of Consciousness Studies, 12*(1), 3–19.

Howard-Snyder, F. (2006). ''Cannot'' implies ''not ought''. *Philosophical Studies, 130*(2), 233–246.

Hume, D. (1748/2000). *An enquiry concerning human understanding*. Oxford: Oxford University Press.

Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv für die Gesamte Physiologie des Menschen und der Tiere, 284*, 1–17.

Laplace, P.-S. (1825/1921). *Essai philosophique sur les probabilities*. Paris: Gauthier-Villars.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences, 8*(4), 529–566.

Libet, B. (1987). Awareness of wanting to move and of moving. *Behavioral and Brain Sciences, 10*(2), 320–321.

Libet, B., Gleason, C. A., Wright, E. W., Jr., & Pearl, D. K. (1983a). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *Brain, 106*, 623–642.

Libet, B., Wright, E. W., Jr., & Gleason, C. A. (1982). Readiness-potentials preceding unrestricted 'spontaneous' vs. pre-planned voluntary acts. *Electroencephalography and Clinical Neurophysiology, 54*, 322–335.

Libet, B., Wright, E. W., Jr., & Gleason, C. A. (1983b). Preparation- or intention-to-act, in relation to pre-event potentials recorded at the vertex. *Electroencephalography and Clinical Neurophysiology, 56*, 367–372.

Nahmias, E. (2005). Agency, authorship, and illusion. *Consciousness and Cognition, 14*, 771–785.

Nietzsche, F. (1886/1973). *Beyond good and evil*. Harmondsworth, UK: Penguin.

Prigogine, I. (2003). *Is future given?*. Singapore: World Scientific.

Sartre, J.-P. (1943). *L'être et le néant*. Paris: Gallimard.

Searle, J. R. (2000). Consciousness, free action and the brain. *Journal of Consciousness Studies, 7*(10), 3–22.

Strawson, G. (2004). Free will. In E. Craig (Ed.), Routledge encyclopedia of philosophy. London: Routledge. Retrieved 8 October, 2005, from http://www.rep.routledge.com/article/V014

van Duijn, M., & Bem, S. (2005). On the alleged illusion of conscious will. *Philosophical Psychology, 180*(6), 699–714.

Wegner, D. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Wegner, D., & Wheatley, N. (1999). Apparent mental causation: Sources of the experience of the will. *American Psychologist, 54*(7), 480–492.

DOI: 10.1002/bsl