## Chapter 1.
# The Ontology of Consciousness

In beginning to examine animacy, consciousness, mental experience, and the like, we feel a certain impotence in having to rely on metaphors rather than on clear objective signs of the phenomena. In this Chapter we will revise and broaden our concept of reality such that mental phenomena are included together with physical objects.

## 1.1. Clarifying the problem

Imagine that an electronics company has created a super-robot to work on Mars. Its appearance does not differ from that of a human being. The robot is supposed to spend two years on Mars, after which its batteries become exhausted and it dies. Imagine next that the robot says publicly that it does not want to fly to Mars, and that the thought of dying causes unbearable suffering. As a result, a mass movement arises in its defense, advocating that the robot be saved from this anguish. The company that created the robot publishes a document claiming that, unlike human beings and animals, robots cannot experience mental pain, and that the robot's "rebellion" is the result of a technical malfunction. The robot is nothing more than an inanimate object such as, for example, a TV set. The robot replies that he is no different from the company's other employees, and that if he does not have mental feelings, they don't have them either, because he can explain every facet of their behavior using only physical concepts. On this basis, he should have the same rights they do.

What is the point of this tale? It highlights the dilemma that the presence or absence of mental experience in an object cannot be diagnosed operationally (Squirs, 1990). In other words, there is no test capable of verifying animacy. Let us emphasize that such a test must differ fundamentally from the famous Turing test, in which a machine and a human being are given the same set of questions: if a human expert cannot distinguish their answers, this means that the machine possesses intellect. Such a test does not allow us to determine whether the machine is endowed with subjectivity.

Many of us believe that any behavior, whether that of a robot or of a human, can be explained within a natural-scientific framework. Looking into ourselves, however, we feel that reducing mental experience to physics eliminates something important. In constructing a natural-scientific picture of ourselves, we leave out our subjective mental domain. The solution may be in constructing a different picture of reality from the one we typically use. I will try to show that Plato's concept of the *ideal* may provide the basis for such a picture.

## 1.2. Plato's ontology

Plato, who lived two and a half thousand years ago, opposed the 'ideal' to the 'material'. This distinction became a cornerstone of Western civilization. It underlies our understanding of ourselves and of the world. This distinction also lies at the core of our understanding of morality (Sayre, 1983).

Unlike us, however, Plato believed that pure ideas exist outside of the human mind and that material objects are only dim reflection of a realm of the ideas. Following Plato, we will call an idea understood in that way the *eidos*.

Plato illustrates the realm of ideas by his famous Myth of the Cave (in *The Republic*). Let us recall this allegory. Shackled prisoners are sitting in an underground cave and cannot turn their heads.

Behind them, there is a crack through which light passes. The prisoners see the shadows of a fence and of various objects that people behind the fence are carrying high above their heads. The prisoners do not know they are seeing shadows: they think it is reality. In the allegory, the objects carried by people are analogues of pure ideas or *eide*. The shadows on the cave wall are analogues of a physical world possible to observe directly. Plato created what might be called "a path to understanding ideal entity." To progress along this path, we have to break the shackles and turn our heads toward the light.

Plato's conception did not remain the dominant one; it was replaced by that of Aristotle, Plato's disciple, who rejected his teacher's scheme concerning the real existence of ideas and stated that ideas exist only as abstractions in human minds. When we call an ideal scheme an abstraction, we follow Aristotle.

Aristotle's conception turned out to be incredibly effective. It led to the development of modern experimental and theoretical science. Despite these successes, however, there remains an unresolved problem. We have made no progress in comprehending the nature of consciousness.

## 1.3. Attempted solutions in physics

Many researchers have tried to find specific physical processes underlying consciousness. Special hopes were raised in connection with quantum mechanics (see Squires, 1990; Barrett, 1999; Satinover, 2001). Unlike the macro-objects studied by classical physics, micro-particles' movements are indeterminate. In the well-known experiment with two slits we can calculate the probability that a particle will hit a certain area on the screen behind the slits, but it is impossible to predict the exact place for any given particle. Since micro-particles' behavior is fundamentally indeterminate, a metaphor appeared in 1930's: the electron has free will. This metaphor is still

used, in one form or another, by researchers who try to explain consciousness on the basis of quantum physics.

Another idea was proposed by Niels Bohr, one of the creators of quantum physics. In the 1920s, Louis de Broglie showed that a micro-particle ought to possess both particular and wave-like features simultaneously. The existence of such a centaur seems to contradict common sense. How can the particle be a body and a wave at the same time? To overcome this problem, Bohr formulated the principle of complementarity: we cannot observe a particle both as a body and as a wave simultaneously. In experiments of one type, we see only a particle; in experiments of another type we see only a wave. This approach underlies the Copenhagen interpretation of quantum mechanics. Bohr suggested using the principle of complementarity as a metaphor for representing mental phenomena (Bohr, 1958).

The most exciting concept in quantum mechanics is the wave function. It sets the distribution of probabilities or of densities of probabilities for various outcomes of experiments. Many researchers have tried and continue to try to understand consciousness as the special realization of a process analogous to the one whose description is given by the wave function. This route seems very attractive, but there is a serious obstacle: in the temperature range suitable for living organisms, the wave function describes only micro-processes. Macro- processes can be described by quantum mechanics only near absolute zero (Penrose, 1989). Consciousness does not seem to be a phenomenon connected to a single micro-particle, but rather with macro-processes of some kind.

In this study, we will elaborate the idea of a mechanism generating distributions of probabilities; we will connect it not only with quantum mechanical phenomena, but also with dynamic systems at their points of chaotic behavior.

## 1.4. The *eidos*-navigator

Let us go back to Plato's ontology and look at the world from the point of view of theoretical mechanics, attempting to find a lacuna within which consciousness may be located. Consider a moving asteroid, a material body interacting gravitationally with many other bodies. We use the term "material body" for a system in which connections between its own elements are stronger than their connections with elements external to the system. The theory of dynamic systems says that some trajectories of a body in movement may be unstable, meaning that the slightest impact is enough for the body to change its trajectory. It is assumed that the change of the trajectory under conditions of bifurcation or more complex ramifications depends on some vanishingly small force not analyzed in the theory itself (see, for example, Peitgen et al., 1992). Now, let us construct an argument making sense of this "vanishingly small impact." Let us imagine that a physical body, by its nature, is equipped with an "ideal navigator" which determines the body's behavior at the points of instability: the navigator generates a distribution of probabilities and transfers the body into a new state according to this distribution.

For the asteroid in the example above, several available trajectories, none of which contradicts the law of universal gravitation, appear at the point of instability, and we suppose that the navigator sets the distribution of probabilities and makes the *choice* of a trajectory requiring an infinitely small impact. In this argument, the material body obeys both the laws of nature and its navigator.

Suppose that the navigator participates in a physical process (not necessarily a mechanical one), if there appear special points at which the process becomes indeterminate. Suppose further that the navigator is a peculiar "factory" producing distributions of probabilities and "infinitely small" impacts directing the evolution of

a material body. This "factory" is equipped with ideal physical machines consuming "ideal" energy. It is important to emphasize that the ideal machines do not consume real energy. To define the relation between the set of material bodies and the set of navigators, we assume that each body has exactly one navigator.

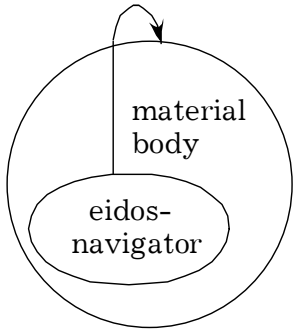


Fig. 1.4.1. A material body and its *eidos*-navigator

What is the correspondence between our scheme and Plato's? Navigators are the ideal entities, *eide*, and the real physical processes are shadows cast by the *eide* on the cave walls. We have two kinds of reality: navigators and material bodies.

Let us take the next step and assume that *a navigator's functioning is the consciousness of the material body to which it corresponds*. Therefore,

> *A material body is animate or living, if it has at least one special point (for example, a bifurcation point) at which its behavior is not determined unambiguously.*

We believe, however, that the degree of animacy depends on the complexity of the physical process. For example, a living organism comprises many complicated processes with a great number of instability points, so that we say it has a complex navigator, but a stone on the ground is not animate, since has neither "special point"

nor a navigator. In the framework of our scheme, the following statements hold:

*A human being has a material body.*
*This body has an eidos-navigator.*
*An ideal physical process is taking place in the navigator.*
*This process is human consciousness.*

Therefore, human behavior is determined not only by the brain, but also by a navigator. If our assumptions are correct, consciousness possesses the structure and functional organization of a certain physical process, that is to say, consciousness is the "form of existence" of an ideal physical process.

## 1.5. Getting inside the navigator

How can we observe the ideal physical process going on inside the navigator? We assume that the consciousness, or, in more general terms, animacy, is an ideal physical process. If, then, we look at a mental process, we can find in it the structure and dynamics of a physical process obeying the fundamental laws of nature, for example, the first and second laws of thermodynamics. Although contemporary psychology does not have methods for detecting mental processes *per se*, in certain cases it may be possible to do.

The general scheme of our work is as follows. We take a psychological phenomenon having both mental and behavioral components, construct a corresponding mathematical model, and conduct empirical and experimental testing of the model. If the model passes the tests, we search for an ideal physical process, some aspects of which are described by the same mathematical model. This physical process (if we succeed in finding it) is a *hypothesis* about what is going on in the navigator. We have to test the hypothesis as well. The test consists of looking for other psychological phenomena

different from the initial ones which can be modeled by other aspects of the ideal physical process so obtained. If such mental processes exist, it means that our hypothesis about the navigator's functioning is not rejected.

Let us emphasize that the physical process has many aspects; it is "richer" than the initial mathematical model of the psychological process. Only some aspects of the physical process are described by the initial mathematical model; there are other aspects described by other mathematical models. Thus, by constructing one mathematical model and finding a hypothetical physical process, we obtain a group of mathematical models and have to test each one by linking them with psychological processes.

Therefore, on the basis of mathematical models of certain psychological phenomena, we arrive at a hypothesis describing an ideal physical process that corresponds to them; we receive a set of other mathematical models and check whether there are psychological processes corresponding to these as well. Underlying this procedure is the expectation that the ideal process in the navigator correlates to many different psychological phenomena. We will begin our work by constructing a mathematical model of bipolar choice.