euresis
Association for the Promotion of the Scientific Endeavour

Volume 3
Summer 2012

# euresis journal

Academic Partners

Fondazione CEUR

Nova Universitas
SCUOLA PER L'ALTA FORMAZIONE

euresis
Association for the Promotion of the Scientific Endeavour

**Cover image**

Magnetic resonance image of the human brain showing colour-coded regions activated by smell stimulus.

**Editors**

Ulisses Barres de Almeida
*Max-Planck-Institut fuer Physik*
*ulisses@mppmu.mpg.de*

Juan Rojo
*TH Unit, PH Division, CERN*
*juan.rojo@cern.ch*

journal@euresis.org

**Academic Partners**
Fondazione CEUR
Consortium Nova Universitas



Euresis Journal (ISSN 2239-2742), a publication of Associazione Euresis, an Association for the Promotion of Scientific Endevour, Via Caduti di Marcinelle 2, 20134 Milano, Italia.
www.euresisjournal.org

Contact information:

Email. info@euresis.org
Tel.+39-022-1085-2225
Fax. +39-022-1085-2222

**Graphic design and layout**
Lorenzo Morabito

**Technical Editor**
Davide PJ Caironi

This document was created using LaTeX 2ε and XeLaTeX

euresis journal

euresis
Association for the Promotion of the Scientific Endeavour

# Letter from the Editors

*Dear reader,*

*with this new issue we reach the third volume of Euresis Journal, an editorial adventure started one year ago with the scope of opening up a novel space of debate and encounter within the scientific and academic communities. From the very beginning, our goals were to propose to a wider public the experience of dialogue we lived in a series of Symposia organized by the Euresis Association, in San Marino, since 2006. The present volume is dedicated to the proceedings of the San Marino 2011 Symposium, centered around the exciting theme of "Brain, mind and language: The mystery of the unity of the self".*

*It is uncommon that a group of renowned scientists and scholars meet up to discuss, not only about the specific technical issues of their disciplines, but also to reflect upon their personal experience of research and the influence that their own discoveries and scientific questionings have on the way they see the world. Of course, this personal dimension of the scientist's experience is one we all share, but it is seldom that we find it expressed and debated with such degree of reflection and seriousness, as has been always witnessed at the San Marino Symposia. Another aspect of richness of these meetings is their interdisciplinary character, meant to follow the universality of the human experience in front of the most pressing questions one can face in scientific and humanistic research.*

*In the 2011 Symposium the theme of debate was the blossoming discipline of neuroscience. More specifically, the complex and little-understood interplay between the brain, the mind and the fundamental functions of consciousness and language. The scientists present at San Marino were invited to reflect upon the results of their own research, and to confront it with the ultimate experience of unity among these concepts that define human beings and characterizes their elementary, albeit mysterious, perception of the self.*

*How is neuroscience contributing, or how can it contribute, to our understanding of these core aspects of human nature? In this volume we are proposing some of the contributions presented at San Marino, which approach this question from multiple angles. Neuroscience is progressively identifying, through the objective approach of the scientific method, the complex networks underlying sensorial, emotional and cognitive functions. But as a discipline, it finds itself in a special position, as all these are capabilities to which we have first-hand access by subjective experience. Neuroscience is not only a fascinating and complex field, but a paradigmatic point of encounter and debate for modern science, with implications to our large conception of what is knowledge, what is the place of science in society, and what is the place of the person in the conception of the world that science helps to build in our days.*

*We hope that the rich and inspired contributions presented in this volume might not only be an enjoyable experience for the reader, but the starting point for a dialogue, and for our own personal journey into these fundamental questions. Once again, enjoy the reading!*

*The Editors.*

# Table of contents


Letter from the Editor pag. 3

Brain, mind and language:
The mystery of the unity of the self pag. 5

*Tommaso Bellini & Marco Bersanelli*
*University of Milan, Italy*

Understanding consciousness:
Need for a sound and reasonable
starting point pag. 9

*Mauro Ceroni & Faustino Savoldi*
*University of Pavia, Italy*

The presence of infinity: Linguistic universals pag. 25

*Andrea Moro*
*University of Pavia, Italy*

Is consciousness epiphenomenal?
Social neuroscience and the case for interacting brains pag. 31

*John T. Cacioppo & Gary G. Berntson*
*University of Chicago, USA*

Can we trust others in our pursuit of knowledge? Anthropological and theological reflections on belief and the neurosciences pag. 51

*Javier Prades*
*San Damaso University, Spain*

euresis journal

euresis
Association for the Promotion of the Scientific Endeavour

Volume 3
Summer 2012

journal homepage
www.euresisjournal.org

# Brain, mind and language: The mystery of the unity of the self

Tommaso Bellini & Marco Bersanelli

Department of Medical Biotechnology, University of Milan, 20122 Milan, Italy
Department of Physics, University of Milan, 20133 Milan, Italy.
tommaso.bellini@unimi.it, marco.bersanelli@unimi.it

Ever since humans have understood that cognitive capacities depend on brain activities, the question of how the human mind relates to the structure and functions of the brain has represented a major intellectual challenge. Indeed, while this topic has a central specific relevance for neuroscience, it is also of great interest for science and philosophy in general, as it directly affects our concept of knowledge and the understanding of our own nature as human beings.

The study of the connection between mind and brain brings with itself the unique challenge of exploring the unity of two apparently unrelated elements, perhaps irreducible to each other: the material structure of our brain and the immaterial reality of our thoughts, mind, self-consciousness. Thus, the brain-mind investigation deals with the ineffable contact between levels of knowledge that are hard to match: the first-person experiential knowledge and the experimental-deductive scientific knowledge.

We all have a direct, evidence-based knowledge of our mind that includes some natural understanding of how it works. Such first-person experience has distinct existential characters that are at the basis of the very concept of individuality in all human cultures: freedom, consciousness, unity, knowledge, language. At the same time, the progress in neuroscience is constantly shedding new light on the functioning of the brain, and it progressively unveils the mechanisms involved in our perceptions and information processes. As this knowledge involves the basic characters of our self, we are often brought to face paradoxical questions, in which the distinct natures of our inner experience and of our scientific knowledge are forced to merge. Is our first-person experience of our mind just an illusion, while the "real" nature of our self is given by the neural circuitry? Or perhaps, vice versa, our scientific knowledge is there precisely because of freedom, consciousness and unity of our inner self? Or should we consider the biophysical structure of our brain as instrumental to our self, with its experience of freedom, unity, and consciousness?

It is quite tempting to ask which is "more real", or "more fundamental", between our physical brain or our immaterial consciousness. Perhaps more humbly, at the current stage of knowl-

edge, we feel that none of these two elements should be denied, and that their apparently illogical coexistence in our nature should be simply recognized – also opening to the possibility that a similar condition might be present in other living creatures as well. Whether all elements could merge into a single comprehensive understanding is a formidable and fascinating challenge.

The Euresis Symposium held in the Republic of San Marino at the end of August 2011 focused on these questions in a friendly, intense and passionate debate, involving scholars from various disciplines. This issue of Euresis Journal reports some of the contributions offered during the Symposium, from which the cultural relevance of the issues at stake, well beyond the boundaries of the scientific progress of neuroscience, is immediately clear. While not attempting a summary of the discussion, we take advantage of these introductory notes to outline some of the issues that have emerged in the debate.

The experience of freedom is at the core of our intimate perception of our self. At the same time, various studies are showing that our freedom is subject to rules, regularities, conditioning. The combination of these evidences provides a powerful incentive to better ponder what we mean by "freedom". A purely conceptual definition would identify freedom with the capability of making a "free choice", i.e., the capacity of unconditionally adopt action A vs. action B as the person is in a given condition. Accordingly, when put repeatedly in the same *identical* situation, the person could act differently. However, since verifying this concept would also require creating exactly identical mind states, this notion of freedom appears as ill-defined. Moreover, the very bases of this approach imply a statistical account of the propensity toward A or B, making it fundamentally contradictory with the first-person experience of freedom, which is not that of a "choice within statistical boundaries". An alternative is to identify freedom in what stands "before" or "outside" the choice, i.e., in the capacity of a person to connect to a notion (programmed action, concept, emotion) involving his/her interest, affection, propensity to act. In this way, freedom would lose its mechanistic role of "decider between options", and would gain the role of polarizing our state when confronted with upcoming possibilities. Indeed, freedom could be better described as the capacity of affectively and rationally adhere to a possibility before acting, rather than the capacity to chose and control the decided action.

A parallel discussion can be formed around consciousness. Again, since each of our conscious states correspond to a brain function, one is easily tempted to identify the two. However, an important distinction has to be made between the functional description of consciousness – sorting sensorial inputs, managing emotions, making decisions, defining concepts – and the qualia, giving the internal qualities of perception, such as "the redness of red", the painfulness of pain, the cheerfulness of laughing, and so on. Functions can in some way be identified with some neural network acting in the brain. Despite the fact that such identifications are often quite indefinite and speculative, knowledge in specific areas of neuroscience

induces us to believe that the functional behavior of the brain will eventually be understood. *Qualia*, however, that provide the texture to our intimate experience, are hardly understood in functional terms. This irreducibility of levels has given strength to dual views that, with widely different expressions, have characterized the western civilization. If and how it may be possible to reconcile these poles of our being into a single view is a matter of discussion and has been a topic of debate at the 2011 San Marino Symposium.

Language appears as a unique crossroad in the investigation on mind and brain. Indeed, language is at the very heart of research on human nature for historical and epistemological reasons. First of all, speculations on the human mind started with reflections on language and the central role of linguistics has never been dismissed up until the so-called analytical philosophy. Second, the structure of our language appears to be unique among all other species and displays clear biologically driven traits. Third, it is so deeply entangled with human reasoning and cognition that the latter domains cannot be approached without exploring the nature and structure of language. The capacity of modulating our linguistic expressions into a seemingly infinite amount of possibilities is a defining property of human experience, and one that still defies algorithmic approaches to language.

Knowledge is another challenging topic, again strongly connected to the duality discussed before. On the one side, we experience the sense of knowing, such as the experience of certainty in cognitive domains, including prototypically mathematics but also moral certainty. On the other hand, knowledge can be gauged on its functional success, witnessed by all the achievements in human history. The objective knowledge and the feeling of knowing appear to match to a surprisingly high degree, investing areas where evolutionary pressure may not have had a significant importance, such as the development of mathematics. To what extent this coincidence goes, is an outstanding question.

Ultimately, all these challenging topics are captured in the – still unexplained and apparently paradoxical – contrast between the unity of our experience of ourselves and of reality, with the plurality of specialized brain functions acting during any of our physical, cognitive or emotional experience.

*Volume 3*
*Summer 2012*

euresis
Association for the Promotion of the Scientific Endeavour

euresis journal

euresis
Association for the Promotion of the Scientific Endeavour

Volume 3
Summer 2012

journal homepage
www.euresisjournal.org

# Understanding consciousness: Need for a sound and reasonable starting point

Mauro Ceroni & Faustino Savoldi

Department of Neurological Sciences, University of Pavia, 27100 Pavia, Italy
mauro.ceroni@unipv.it

## Abstract

*When I began to analyze the starting point of various authors in the field of neuroscience, I realized that this has already been done by others who have dedicated themselves to it for a long time and who have done so certainly better than I could do it. Therefore here I will but draw from these authors who, in my opinion, have made this issue clearer. So many answers have been already given, but they are submerged in an ocean of repetitions, ambiguity and partial truths. They are neither underlined, nor put in the right evidence. What is most lacking seems to be a love for the truth which is stronger than our prejudice. At the end of this paper, I shall report on my own personal position as a provocation to the discussion.*

## 1. The problem of consciousness

In the history of Western thought, one of the most representative philosophical movements identified with the problem of consciousness, understood and approached it as "the consciousness that man has of himself", that is, the main road leading to the complete and certain understanding of man. This trend began with Plato when in the *Phaedo* he affirms that the soul grasps the truth only when it is independent from the senses. This topic is taken up by Plotinus: "You can see the wisdom and the justice without leaving the soul; the soul sees them in itself, reflecting on itself." [1] The theme of self-reflection is fundamental, especially in the work of St. Augustine, who applies the method in a radical and coherent way. Self-reflection allows us to grasp in our inner self "a presence that is deeper than ourselves". This attitude will mark the Western philosophy and theology until the advent of Christian and atheist existentialism, Husserl's phenomenology, and many currents of modern spiritualism. *"Noli foras ire, in te ipsum redi, in interiore homine habitat veritas"* – "Do not go outside, return into yourself, in the inner man dwells the truth" [2]

Edmund Husserl closes his *Cartesian Meditations* by quoting this last sentence of Augustine and, in all his strenuous efforts, he remains faithful to Augustine's exhortation. In fact, Husserl called his meditations "Cartesian", since he saw in Descartes the resurgence of Augustinian themes. The Cartesian *cogito* is the immediate, clear and infallible certainty the ego has of his inner life in *interiore homine*.

In line with Descartes' thought is Sartre, the most "Cartesian" among the modern philosophers, who maintains that "...the consciousness of being is the being of consciousness..." [11]; or "For it existing and being conscious of existing is one thing". In other words, the great ontological law of consciousness is the following: the only mode of existence of consciousness is to be conscious that it exists [11], which means that the specificity of human existence is the consciousness to exist.

The scientific psychology of the turn of the '800 to '900 (with Franz Brentano, William James, Wilhelm Wundt) believed that consciousness was the central problem of scientific psychology and, consequently, it had to be tackled as such. Wundt and his school tried to develop an experimental methodology that would allow the scientific study of consciousness: this method was based on introspection. Many psychologists of the time believed that the royal road that leads to consciousness was that of introspection, but introspection, despite the efforts of Wundt's school, still remained somehow subjective and, as such, could not meet the criteria of scientific objectivity. The reaction to this concept still based on introspection led, at the passage from the XVIIIth to the XIXth century, to the birth of behaviorism.

## 2. Artificial intelligence and philosophy (H. Dreyfus)

The advent of computers has fueled the discussion among the authors who study the philosophy of mind. According to Thomas Nagel, the introduction of computers is part of the reductionist programme that dominates the current work in philosophy of mind [4]. This programme is completely misleading and is based on the assumption that a particular conception of objective reality exhausts the reality itself [4].

> *In the end, I think, it will be evident that the ongoing efforts, in order to understand the mind by analogy with computers that can perform superbly some of the same outside tasks of conscious beings, constitute enormous waste of time. The real principles underlying the mind will be discovered, if they ever will, only by a more direct approach.*

The most radical philosophical critique of computer science and artificial intelligence is Hubert Dreyfus's. His main work, *What computers can not do: The limits of artificial intelligence*, dates back to 1972 (II ed. 1979) and still retains its validity. There, Dreyfus performs a critical analysis of the theoretical foundations of computer science. According to Dreyfus,

these fundamentals are derived from a philosophical tradition of the Western world that goes from Plato to the neo-empiricists; he moves from philosophical positions inspired by the philosophy of Heidegger and Merleau-Ponty.

The assumption that man works as a computer, that is, as a facility that processes symbols, implies the following postulates:

1. *organic postulate:* according to which the neurons process the information according to discrete steps, using the biological equivalent of the process on/off;
2. *psychological postulate:* according to which the mind is seen as a device that works on bits of information according to formal rules. Empiricists and idealists would, according to Dreyfus, have set the stage for this thought pattern, defined as data processing, a third-person process in which the elaborator's involvement is not an essential part;
3. *epistemological postulate:* which states that all knowledge can be formalized, that is to say that everything that can be understood can be expressed in terms of a logical relation, more exactly, in terms of the Boolean function, the logical calculus that governs the ways in which the bits are related according to rules;
4. *ontological postulate:* as any information you enter the computer must be in bits, the computer - model of the mind - assumes that all relevant information concerning the world, all things essential to the production of intelligent behaviour must be analyzed in principle as a set of specific elements independent of the situation. This is an ontological presupposition according to which what exists is a series of events logically independent of each other.

The four conditions are considered by scholars of artificial intelligence and cognitive simulation as axioms that guarantee the results, while they are in fact only possible hypotheses. None of the four postulates is justified on the basis of empirical or theoretical arguments.

Dreyfus debates and criticizes all the postulates and demonstrates that these are derived from a conception of the mind that is intended as a device capable of calculating, based on clear rules, according to a sequence of distinct steps, a series of data that he defines atomic and neutral. This view, Dreyfus says, is a tidal wave produced by the confluence of two powerful currents:

- the Platonic reduction of reasoning to explicit-rules and the reduction of the world to atomic facts to which these rules can be applied without risk of interpretations;
- the invention of the digital computer, an information processing device that formulates calculations in accordance with explicit rules and evaluates data in terms of logical elements independent of each other. In our culture, the computer seems to be the paradigm of logical intelligence.

There are also three areas necessarily neglected by scholars of cognitive simulation and artificial intelligence which are necessary to explore as underlying all intelligent behavior. These areas include:

1. the role of the body to organize and unify our experiences of objects;
2. the role of the situation that provides a backdrop against which our behaviour can be ordered without rules;
3. the role of the goals and needs.

Those who accept the epistemological premise, according to which human behaviour must be transcribed into the formal language of a heuristic program for a computer, have to develop a theory of intelligent behaviour that will not resort to the fact that man has a body, because the computer does not have the body. In thinking that we can do without the body, these authors follow the Western tradition, ranging from Plato to Descartes, which states that the body is not at all essential to intelligence and reason. If the body appeared essential to intelligent behaviour, we should wonder whether the body can be simulated in a computer program. If this is not possible, the project of artificial intelligence is doomed to failure from the start. Computer technology can successfully deal with the ideal languages and with abstract logical relations[1]. What computers exclude is that sort of intelligence we share with animals and that has survived the simulations of machines. When the human mind recognizes objects in space or time, it does not proceed enumerating a list of isolable, neutral specific features. In other words, the mind does not proceed from atomic elements to totality, but grasping the parts in a whole. The notes of a melody have value because they are perceived as part of a melodic series and not vice versa, and the same applies to the elements of a sentence. The meaning of the details is determined by our perception of the whole: these are the teachings of *Gestalt* psychology and phenomenological philosophy.

In conclusion, the pattern recognition is relatively easy for a computer if the pattern is defined by few specific traits, but in the case of complex models, the computer does not work. According to the phenomenology of Merleau-Ponty, humans recognize complex patterns thanks to a capacity that is actively and organically linked with the body that responds to the environment by virtue of the continuing sense of its operation and its goals.

At this point, Dreyfus introduces a fundamental concept of the philosophy of existence: the situation. According to this concept, every thought, every action is understood from a specific human commitment with the world and history.

[1]It should also be pointed out that computers, however sophisticated, do not understand the concept of infinity as all their operations always give finite numbers.

The situation or context is the human mode of being in the world and the situation makes it possible to conduct orderly, but not subjected to formal rules. The open-structure problems, unlike the games and tests, have three levels of difficulty:

- they determine what facts are possibly relevant;
- what facts are actually relevant;
- among them, which are essential and which are not essential.

First, in a given situation, not all facts are possibly relevant: some are, but most of them are irrelevant. Since the computer is not in the situation, it must deal, at a time, with all facts as possibly relevant. Dreyfus states that we are at home in the world and we can find our way because it is our world, produced by us as context by our practical activity. The world or situation allows us to aim at the meaningful objects it contains. As a field of experience, it is structured by our tasks and is linked to our goals that, in turn, correspond to our social and individual needs that have created the world.

The world is the place of history and in history cultural revolutions are carried out (the Greek civilization, Christianity) that change not only the visions of the world, but also what Kierkegaard calls the spheres of existence; in addition, the changes also involve the conceptual level.

The cultural revolutions show us, as Pascal had first realized, that the boundaries between nature and culture are not clearly defined. Instinctual needs can also be modified and even the paradigms may change; therefore, human nature is not fixed forever; human nature is very malleable and could also be about to change. If the computer paradigm becomes so strong that men begin to think of themselves as if they were digital devices, made on the model of artificial intelligence machines, then, since the machines can not be like men on the grounds that we have shown, humans can gradually become like machines: "The risk is not the advent of the superintelligent computer, but of intellectually underdeveloped human beings" [5].

If we reflect on the mind of man, we can not exclude the aspect of always being in a context, and, therefore, experience is already in situation, in a situation where the facts have already been interpreted. This theory also suggests the final state in which human beings are: it depends on their projects, which in turn are a function of their body and their needs and these needs are not fixed once and for all, but are interpreted and become crucial because of the cultural background and the changes of human self-interpretation. Then we can understand why there are not facts made with a built-in meaning, or fixed human forms of life that can be formalized.

In other words, computers are not in a context and do not have a body, while the intelligence of human beings is always in a context and is conditioned by the fact that man has a body, "What distinguishes men from computer, no matter how cleverly they are designed, is not an abstract, universal, immaterial soul, but a concrete, specific, material body" [5]. Human intelligence is always in situation and this implies an original background of beliefs, namely, common sense. These beliefs are not objectively measurable, and therefore can not be formalized or simulated; the intelligibility and the intelligent behaviour must be related to the common sense of what we are, this means necessarily a kind of knowledge that can not be made explicit if we want to avoid the infinite regress.

## 3. McGinn's controversies

According to McGinn, it is only in the 90's that the problem of consciousness reappears on the scene. Then, the philosophers admit that the philosophical problem of consciousness is real and not the result of mental confusion, while the neuroscientits are beginning to build neural-mind connections and recognize that the brain is the seat of consciousness. At this point, there is a problem: once admitted that consciousness is a peculiar phenomenon in the natural world, we must find a place for it in our scheme of things and give an explanation of its nature. How does consciousness fit into the scientific world so laboriously built in the seventeenth century? How is it related to a physical world consisting of atoms, space, force fields? How can the brain bring consciousness into existence? This raises disturbing issues, once you stop to deny the existence of consciousness. But there was a reason for this denial: consciousness is threatening. It appears as an anomaly in our conception of the universe; it is a site where our mental models of understanding are no longer valid. How - Colin McGinn wonders - can an objective science of nature, which studies the particles and their modes of aggregation, find a place for the subjective phenomenon of consciousness? How can some brain cells build the experience of seeing red or the emotions of despair? May consciousness exist out of the accessible world of the natural sciences? May the dualism of body and mind, long rejected, be the correct position? Has the ghost that haunts the machine returned? Or worse, is perhaps the car a masked ghost? May consciousness be phantomly rooted in the matter or is matter, perhaps, less material than we think? McGinn believes that the renewed interest in consciousness represents the next major phase of human thought.

We now begin to deal with an aspect of nature we do not understand. It is unclear whether our efforts will be successful. Obviously, there is no unanimous consensus on the issue of consciousness, on the contrary, there is a radical difference of opinions. McGinn illustrates this difference in referring to two books: John Searle's *'Mind, Language and Society'* and Paul M. Churchland and Patricia Smith Churchland's *'On The Contrary'*.

Searle believes that consciousness is fundamentally irreducible to the terms of neuroscience;

according to Searle, neurophysiology as a whole is unable to provide an adequate explanation of the true nature of consciousness, although the neuronal processes form the basis of conscious activity. Searle also believes that consciousness is the fundamental problem of science and the philosophy of mind. The Churchlands sway, however, between negationism (eliminativism) and the claim that consciousness is completely reducible to neurophysiology. The difference between the two positions is very large, comparable to that which divides the darwinists and the creationists. For his part, McGinn does not share any of the two positions, particularly with regard to their foundations.

Searle defines its own solution to the problem of mind/body with the term "organic naturalism" and summarizes his position in these words: "consciousness is caused by brain processes and is the highest manifestation of the brain system." [6] The idea is that third-person phenomena of the brain (neurons and their activities) operate to cause high-level subjective processes that have what Searle calls first-person ontology, i.e. processes that exist as they are experienced by a conscious person or subject. It is a fact of neurobiology that certain brain processes can lead to states of consciousness: we should accept these facts without the metaphysical implication that traditionally marks them.

According to McGinn, this conception is deceptively simple. The conscious processes are different from the standard physical processes taking place in the brain and are defined by their subjectivity. They are also biological processes for three features. These processes (1) occur in an organic system, unlike computer programmes, (2) result from processes of natural relationship and not from an intentional design, and (3) are genetically based rather than learned or acquired, as opposed to the knowledge of history or to typical performances in certain skills. In addition, brain alteration modifies states of consciousness and brain function is a precondition for the existence of states of consciousness. Finally, states of consciousness are the high level properties, in the sense that "they do not belong to the isolated primary components of the brain, but result from a combination of these elements into a complex organism." [7] The question - McGinn wonders - is as follows: "Is really this the solution to the mind/body problem?" Searle in fact likens consciousness to other high-level macrophenomena that are composed of lower-level micro-phenomena, such as solid and liquid states, digestion, photosynthesis, and these phenomena are, according to him, fully explicable on the basis of micro-processes that underlie them. For McGin, consciousness, however is something that goes beyond the neurophysiological facts which determine it. Consciousness is not reducible to its physiologic/causal bases.

While it is easy to understand the relationship between the state of the molecules and certain physical states, in the case of consciousness, however, we are faced with a mode of inexplicable dependence, unique in nature: the dependence of subjective facts from objective facts. But how can this happen? Suppose that the visual experience of red is caused by a number of neurons that discharge in the occipital cortex. The question is this: how can a subjective

experience depend on the activity of simple electrochemical cells? What a cell has to do with experience? Searle does not say anything about the concept of supervenience, and the omission is crucial. The notion of supervenience implies that the conscious mental state of a subject is completely determined by physical states of the brain. But another problem immediately arises: what is determined in accordance with the supervenience? What's in the neurons, that enables them to determine consciousness? Searle has no answer to the central question, or rather he does not even raise the problem, but he merely makes statements.

In fact, Searle asserts that it is a fact of nature that consciousness is produced in this way, but does not explain how the subjective consciousness may result from transactions carried out by a small (sic!) number of gray cells grouped together. Searle might respond to these objections by saying that it is a purely scientific problem and not a philosophical or metaphysical one. And in fact he has done all the philosophical work when he developed his own theses – it is now up to the empirical science to discover the actual mode of dependency that links consciousness to the brain. But a similar response has no value: in fact, it does not really matter whether we label the problem "scientific" or "philosophical", the fact remains that a theoretical problem is still unsolved, "a problem we do not have the slightest idea of how can be solved."

We do not know how neurons have the ability to generate a conscious episode. Neurophysiologists find correlations between brain states and states of consciousness, but nothing in neurophysiology begins to explain this correlation. There is not even an explanation why organisms with brains are capable of sensations or feelings. Philosophers are interested in this problem as opposed to the mechanical problem of how to derive the liquid state by water molecules, as the problem related to consciousness is a conceptual one, because it seems to test our conception of the mind and the brain.

The concepts of consciousness and brain seem inherently inadequate to allow an explanatory theory linking them together. Consciousness is not an observable phenomenon at all. We can not see someone's conscious state, even if we looked inside their brain, watching the brain cortex and biochemical reactions, because states of consciousness are not the kind of things that can be observed. The underlying brain processes are apparently observable physical events, and this feature distinguishes consciousness from other high level phenomena and remains once we confidently assert that consciousness is a biological phenomenon of high level. McGinn responds to Searle's assertion saying that this is the beginning of the problem, not the end.

Even Searle's arguments on intentionality are questionable: McGinn properly defines intentionality as the ability of the mind to have objects, to have a meaning or content, to go beyond itself. Intentionality is what makes an animal to be a semantic one. Searle has done much to emphasize the importance of intentionality and said very sensible things on the subject, but

he is a philosopher and has an innate resistance to admit that he is embarrassed. In recent years, many philosophers have tried to naturalize intentionality to make it understandable by reducing it to something more familiar: causation, biological function, computational structure, in which cases the intentionality emerges as nothing, yet, as a special case that we have on the list of the acceptable facts from a scientific perspective.

Searle has not done any of this, but claims to have his own explanation of the nature of intentionality. Instead of reducing intentionality to something else, as other authors do, he declares it irreducible while offering an explanation as to make it biologically natural: as an example he suggests the physiological processes underlying the phenomenon of thirst. A lack of water in the body causes an increase of neuronal discharges in the hypothalamus through certain biochemical mechanisms. In turn, the neuronal discharges induce in animals a conscious desire to drink. Since the desire to drink is an intentional state, i.e. it is directed toward the act of drinking, this provides an explanation of a conscious mode of intentionality. In the same way, other forms of intentionality are developed: perceptual, cognitive, etc. Thus, we manage to explain the intentionality biologically, but this is not a naturalistic explanation of intentionality itself. All this tells us about the physiological mechanisms underlying intentionality. The philosophers interested in understanding intentionality want to realize what an intentional relationship is and what it consists of, what it means for the mind to be directed to something out of itself when we perform the act of thinking or desiring or perceiving.

What is the mysterious relationship of *aboutness* which our true mental states show? What is the nature of mental representation? Searle's book does not account for these problems. He simply describes what causes a mental state that shows intention, and, therefore, leaves the conceptual problem unsolved. How can a brain be able to originate mental states that represent the outside world? What is in certain groups of neurons that makes them capable of symbolic activity? The kidney cells have no intentionality; why, instead, do brain cells have it? What is the relationship between the cells of my brain and the city of London in the event that I am thinking about London?

Searle's conception is not false, but irrelevant, as it is not an answer. It would be better for him to remain true to its claim of irreducibility of intentionality and not try to explain it naturally. The problem remains of how intentionality is possible in a physical system. Searle tries to go both ways: he declares that a conceptually puzzling phenomenon is irreducible, but strives to have something left unexplained.

McGinn criticizes the eliminativistic position of Paul and Patricia Churchland.The arguments in favour of eliminativism are rather weak: first, there is no argument for the falsity of folk psychology that covers everything inside the mind: bias does not imply error. Second, the durability of folk psychology could not be explained by its dogma, but to its obvious

truth: compare in this respect with the stability of elementary arithmetic that lasts from the time of the Greeks. Churchland's third objection is deceitful, since it is assumed that folk psychology is unable to integrate with the natural sciences of man. In fact, the model of contemporary cognitive science shows continuity with the apparatus of popular psychology. The philosopher Jerry Fodor, for example, has convincingly argued that the conception of the mind, understood as an information analyzer and processor capable of manipulating symbols, adapts easily to the framework of the mind, as it is understood by the popular psychology, namely as a coherent range of propositional attitudes such as beliefs and desires [8]. Therefore, none of these arguments shows that folk psychology is radically wrong about the topics that interest us.

On the other hand, Churchland underestimates the first-person aspect of folk psychology. This is not a speculative theory we apply to others, it is the means by which we refer directly to our personal mental states. This immediate first-person reference requires special privileges: the knowledge I have of the fact that, just right now, I am thinking about philosophy is as certain as any form of knowledge can be, it is almost incorrigible, as Descartes thought. But Churchland believes that this form of firsthand knowledge is not knowledge at all, since I have no thoughts, according to the eliminativistic doctrine. Folk psychology is for him only a false theory and not the means of such immediate, absolutely certain, first-person knowledge.

But once we recognize the privileges of the person in folk psychology, it becomes inconceivable that we may be wrong on having mental states in an exclusively privileged way. Hence the belief, well founded, against eliminativism: we just know that we have beliefs and desires, and everything else. One of the recurring themes in Churchland is that neuronal dynamics underlying what we will call the mind does not involve the symbolic representation of a propositional kind. The brain is not constituted as a device to analyze and process the internal language that underlies our cognitive capacities. Instead, he writes that neurons direct themselves into "activation vectors", patterns of activities that do not involve anything resembling a sentence.

For its part, folk psychology insists on describing the thinking using the language of propositions, but Churchland argues that there is nothing in the brain that corresponds to the propositional apparatus of popular psychology. McGinn focuses on two points: (1) we are convinced, unlike Churchland, that folk psychology simply provides a description that ignores the details of what neurons are doing, like the description of a computer software ignores the description of the hardware. From the perspectives of elementary physics, activation vectors do not exist in the brain, which is a collection of subatomic particles. Reality is manifested at different levels and what is invisible to a level may seem obvious to another. (2) Churchland carefully avoids confronting the issue of language which analyzes itself. But if we understand language we must assume that our mental dynamics involve the manipula-

tion of linguistic propositional structures, because language consists of sentences, and if the propositional attitudes such as beliefs and desires are related to the language, they involve the internal linguistic structure.

The representative apparatus of the brain may not have, to a certain level, any reference to linguistic symbols, but it does not follow that these symbols do not play a part in the operation of our mind. Once we admit this, popular psychology can claim a science of the brain. This is precisely the position of Jerry Fodor that points out to "the language of thought".

Churchland's insistence on non-propositional neural representations leads to a strange result: it is more plausible that a computer thinks than humans do. Churchland notes that the standard architecture of computers consists of manipulating propositional structures serially while the brain works through non-propositional neural activations connected in parallel (which are called connectionist networks). Then the computer shows the shape of the internal processes that folk psychology requires, while human brains show a completely different type of mechanism. The result is that eliminationism is closer to the reality of computers than to ours. According to Churchland, the computer I am using has more claims than I have to be a thinker. This is a *reductio ad absurdum* of Churchland's positions.

Churchland and Searle represent opposite poles of the philosophical world. Searle takes seriously the notions of common sense about the mind and resists attempts to reduce or eliminate the mind in favor of a materialist metaphysics. Churchland believes that the idea that humans have beliefs and desires is a false theory about the way our brains work, a theory which must be replaced by a better theory that describes man in accordance with neurosciences. But does a middle way exist?

According to McGinn, the two theories are understandable answers, though incorrect, to a conceptual problem. The problem is how to integrate the conscious mind in the physical brain, how to capture the unity beneath the apparent diversity? The problem is very difficult and McGinn does not believe that someone has any strong ideas on how to solve it. However, we can expect two types of responses: either there is unity with the mind or there is no mind to unify with the brain. Then we get the antireductionist duality of Searle (who does not like this definition) or maybe we will have the combined eliminationism of Churchland, without consistency, with an attempt to reduce mental phenomena to our understanding of the brain. McGinn believes that the best solution is represented by a theory that unifies the conscious mind with the physical brain, but we have no idea what this theory would be. In fact, there is likely an underlying unity, even if we do not understand it.

McGinn affirms that there must be a unity below, because, were there not, we should postulate a miraculous form of emergency in the biological world. Consciousness can not jump into existence from matter as the Aladdin lamp genie. But our ways of access to conscious-

ness and the brain, respectively, through introspection and sensory perception, can not in principle reveal the hidden structure of this indispensable link.

> *I know that I feel pain because I feel pain inside me and I know that my neurons are activated and, using a scientific method, I can observe that they are, but I neither have the immediate consciousness of the necessary bond that unites all sensations and brain processes nor I possess any methods to extrapolate this relationship. We cannot infer our internal brain states from our consciousness and we cannot infer consciousness from the immediate sensory awareness of our brain; so the mode of their association is beyond our cognitive faculties.[7]*

We can know each side of the great fissure that divides the mind and body, but we have no faculty which reveals to us how they are split off from each other. All this is the basis of the difficulties we face in formulating a theory that can connect consciousness with the brain, but it is surprisingly hard to find that not every aspect of the natural world adapts to our powers of comprehension. We do not expect other evolved species to be omniscient. This is because we believe that our intelligence has evolved with the ability to solve any problem concerning the universe of which we are a small contingent part, but even if this argument is wrong we should admit the possibility that our knowledge about the mind and brain is severely limited and this produces the impression that this association is a brute and inexplicable fact. This could be an explanatory theory.

## 4. A personal proposal

It is absolutely clear that every thought, feeling, movement, perception, experience has neurophysiological and electrophysiological correlates. I do not exist separate from my body into the space-time dimension in which we all live. Nothing can happen in me that does not have a neurological basis, which does not involve the activation of neural circuits, but this does not mean that everything is due to my brain, as I'm not just my brain.

So let us ask ourselves: how can we tackle the question of how man is made, what constitutes consciousness, the I? First, this issue can not be resolved by science, it is not a science based issue. When science purports to answer these questions it makes a mistake, from which science is blinded: the claim that the scientific method is the only method of knowledge. When such a claim is made, one implies that what can not be known by the scientific method does not exist. Also, the statement that "scientific knowledge is the only objective true knowledge" is not a scientific statement, but a statement on the philosophical ground and as such should be sustainable.

How do I know the constituent factors of myself? If the criterion for this investigation was external to me I would be alienated, the slave of whoever owns the criterion. If the criterion for the knowledge of the constituent factors of myself was to be strictly scientific, it would be the paradoxical possibility that one other than myself could know myself better than I!

Therefore, the criterion has to be inside me, and consequently, the method for such knowledge must be akin to a look into my personal experience, as reflected in the thick of life, in action.

Within this dynamics I can easily find that I'm composed of two sets of factors, with different characteristics that are irreducible to one another. The first one includes a widespread phenomena in space, divisible, measurable, visible, and which inevitably change over time, mutate, become corrupt (we call this order of phenomena of the material, it is "the materiality"). The second one, such as the concepts, mathematical truths, the Self, value judgments, the decisions we take in life, are not measurable, non-divisible, non-perishable, non-visible; it is an order of phenomena which we can call spiritual, "not material", or in a more restricted way it could be called mental. Therefore, these two orders of phenomena are irreducible to each other and at the same time given only together.

Note that this reading of the experience does not contradict any of the data that the neurosciences have given us. Every action of man always implies a modification of the brain's neurophysiology and electrophysiology (you can not separate the two principles, as body and brain cannot be separated into the space-time dimension in which we live); but it also stresses that any attempt to fully explain the phenomena of non-material from material principle clashes with the experience that gives them to us as fundamentally irreconcilable. Such a position does not mean that they are two juxtaposed realities, it indeed shows that they belong to and form a unitary being, the human person. Therefore, this reading is distinct from the philosophical position that has most influenced the modern world – I speak, of course, of Descartes.

The revolutionary character of Descartes is the introduction of the "first person" to the foundation of philosophical discourse, the *"cogito ergo sum"*.

> *The introduction of the first-person perspective produces a radical renewal of the issue on (the ego, the subject or, in terms of the debate of the times of Descartes) the soul, which becomes consciousness in the eminent sense. [9]*

And yet, this is done with a clear stance with relevant consequences.[2]

> *The thinking substance (the cogito) is the one whose whole essence or nature, is that, in thinking and in order to be, does not require any place nor depends on any material thing [...] This is why the "I", i.e. the soul, what makes me myself, is absolutely distinct from the body and is also easier to recognize than the body, and even if the body were not, the soul does not cease to be what it is. [9]*

For Descartes, therefore, the ego coincides with the *res cogitans* that has existence in itself, and the body, the *res extensa*, is a pure machine, a mechanical device, somehow united to the

[2]From Descartes, Discourse on Method, quoted by Vanzago [9].

self. The dualism is established with its insoluble problem: how the two substances interact?

We know the response given by Descartes: the soul resides in the epiphysis (a centroencefalica median formation) and hence she governs the body. The Nobel Prize for biology, Francis Crick, discoverer of the structure of DNA with Watson, a decade ago proposed that human consciousness is established in the claustrum, a thin gray bandeletta immediately subcortical in the temporo-parietal-frontal lobes, of which we ignore the function. This was and is the influence of Descartes in the history of the western culture!

The main topic of all reductionists, eliminativists, scientists, materialists is that every activity, thought, feeling, desire or experience implies (but not depends upon) brain activity. These data are reported as a novelty made possible by modern exploration techniques of cortical functions. But neurologists knew about this for almost 200 years: the structure and functions of the brain were, in fact, discovered by observing and studying these natural experiments, which are the diseases of the CNS. For example, neurologists know that a lesion of Wernike's cortical area, in the posterior third of the superior temporal gyrus, that allows the understanding of spoken and written language, determines the appearance of a real and severe dementia: the disappearance of the opportunity to realize the meaning of words implies the loss of opportunity to think. Therefore it is important to realize that this does not imply that mental activity is then generated, determined, caused by brain activity!

The challenge before us, in front of Neuroscience (we are only on the initial threshold of understanding brain functioning) seems to me well expressed by the words not of a neuroscientist, but a theologian (remember Theillard de Chardin):

> *Nature - the flesh, bones, viscera, the cells (one might add, especially the brain, neurons) – become in the human beings need of the infinite [...] Unlike animals, even our physiology is whole set with this opening to infinity: this opening to infinity is rooted in our humanity, for this it is ineradicable. (J. Carrón)*

## References

1 Plotinus, *Enneads*, IV, 7, 10.
2 St. Augustine, *De Vera Religione*, par. 39.
3 Sartre, J.-P. 2010, *L'Imagination* – The Imaginary: A Phenomenological Psychology of the Imagination (Routledge Classics)
4 Nagel, T. 1989, The View from Nowhere (OUP USA).
5 Dreyfus, H.L. 1978, What computers cannot do: The limits of artificial intelligence (Harpercollins).
6 Searle, J. 2000, Mind, Language and Society (Phoenix).
7 McGinn, C., 'Can we Ever Understand Consciousness?', New York Rev. of Books, June 10, 1999.

8 Fodor, Ja 1979, The Language of Thought (Harvard University Press: Cambridge, Mass.)
9 Vanzago, L. 2009, *Breve Storia dell'Anima* (Il Mulino).
10 Churchland, P.M. and P.S. Churchland 1999, On the Contrary: Critical Essays 1987-97 (MIT Press: Cambridge, Massachuss.)
11 Sartre, J.-P. 2003, *L'etre e le neant* – Being and Nothingness: An Essay on Phenomenological Ontology (Routledge Classics).

*Volume 3*
*Summer 2012*

euresis
Association for the Promotion of the Scientific Endeavour

euresis journal

euresis
Association for the Promotion of the Scientific Endeavour

Volume 3
Summer 2012

journal homepage
www.euresisjournal.org

# The presence of infinity: Linguistic universals

Andrea Moro

Institute for Advanced Study, IUSS Palazzo del Broletto, 27100 Pavia, Italy
andrea.moro@iusspavia.it

## Abstract

*Language and languages: trying to reconcile these two facts has been maybe one of the most important motivations in the history of linguistics; certainly this is what today represents the most important challenge in understanding the nature of the mind, and at the end, of human beings. At which point are we today on the research of the linguistic universals? Let's take a step back and have a look at this.*

## 1. Introduction

There exists two facts which belong to the basic experience of almost everyone: on the one hand, language appears like an universal experience, elementary, as fundamental and constitutive of the human being as to be able to define him: The human being is the being capable of language. On the other hand, the astonishing recognition of the diversity of languages: the irreducible experience of the particular, of the restrained, of the chaotic. Obviously, there are also marked differences in the expressive capacity of individuals, especially in the number of known words, but this is negligible compared to the substantial sharing of the immense complexity of the linguistic code by part of all human beings. We can consider for example the system of verbs declination or the almost unsurpassable difficulty of giving explicit meaning to words of common use, like the very simple case of the word "maybe". It is not, that is for sure, a strange word, but to define its meaning is an extremely complex enterprise that requires calculations and sophisticated formal models. Language and languages: trying to reconcile these two facts has been maybe one of the most important motivations in the history of linguistics; certainly this is what today represents the most important challenge in understanding the nature of the mind, and at the end, of human beings. At which point are we today on the research of the linguistic universals? Let's take a step back.

## 2. Universality of the language forms

Roger Bacon, the franciscan monk known by his contemporaries as *"Doctor Mirabilis"*, one of the greatest philosophers of the middle ages, summarized the idea of the universality of language forms in an unequivocal way: "Grammar is one and only one according to its substance in all languages, even if there can be accidental variations." This conclusion, literally deduced from the hypothesis — guaranteed from the theological side –- of a substantial symmetry between perception, language and reality, could not contrast in a neater way with that of Martin Joos, an American linguist, that correctly summarizes the dominant convictions of the middle of last century: "Languages can differ one from another without limits and in an unpredictable way." This was also an ideological deduction, so to speak; that is, substantially based on a theoretical prejudice, namely, that a language is, in all of its aspects, a purely arbitrary convention.

This chaotic vision has been recently proved false, both on the formal side [1, 2, 3] and in the neuropsychological side (see [4] and references therein). Nevertheless, apart from the *lack of* experimental studies that rendered it plausible, it is interesting to note how this vision was welcomed because of the defence of epistemological relativism that it carried associated to it, as it was perfectly suited to justify a technological vision of the mind that today seems to return dressed as "biological discovery." In addition, the effort towards the reduction of cognitive capacities to formal mechanisms that are essentially predictable once the contextual conditions have been defined –- an idea once called "cybernetics", and that now resists, even with a loss of popularity, with the tag of "artificial intelligence" — was also sustained by a mobilization of funds and people that in fact constituted a mode to recycle the experiences accumulated in the sector of military communications during the second World War.

There is also in this case a direct testimonial that I believe will be clearer than any other elucubration. Who speaks are logicians and mathematicians of great fame, from the electronics laboratory of one of the most prestigious technical universities of the United States, the Massachusetts Institute of Technology: "There was in the laboratory the general and irresistible idea that with the new knowledge of cybernetics and with the recent techniques in information theory we had arrived to the last step towards the complete understanding of the complexity of communications between animals and between machines" [5]. It was precisely at the MIT that, also as a reaction to this reductionism, Noam Chomsky showed, using a mathematical model, that none of the known algorithms could automatically generate a complex structure like that of human language [6].

With this, Chomsky recognizes immediately that the heart of human languages is constituted by a capacity of manipulating primordial elements (words) producing potentially in-

finite structures ( phrases) following schemes that are discovered just as physical laws are discovered, traditionally known as syntax. The manifestation of infinity on a basis of finite elements — i.e., syntax – can be qualified as the distinctive tract of all human languages, and therefore of language as well.

This discovery has in fact completely changed, not only the status of linguistics, but also that of neuroscience in general, putting the language back at the centre of empirical studies and in many cases making language the model for the study of other cognitive capacities, like those related to mathematics and to music. There exists at least three important consequences that follow form this first intuition.

The first consequence follows directly from Chomsky and can be immediately understood from this quote: "The fact that all normal children acquire comparable grammars of great complexity with notable ease suggests that human beings are *in some mode designed* for this activity, with a capacity of dealing with data and to formulate hypothesis of unknown nature and complexity" [7]. The second consequence is in some sense implicit in the first one: if man is designed in an special way, this design must be somehow established biologically, and thus it should be possible to trace back the neurobiological elements to which they are correlated; such elements cannot but be universal, as universal are all biological features of human beings. Such an intuition, based on observational data of a comparative kind, has been verified in a substantial way in the last decade by radically innovative experiments performed using the techniques of neuroimaging. The clinical foundation, that since always had constituted the master way for the study of the biological fundamentals of language (see for example the classic work of Lennerberg [8]), is indeed now complemented by new methods that avoid the need of proceeding only in the presence of pathologies.

## 3. The limits of Babel

The universal linguistics, at least that related to the syntax, can be in some way traced back to the functional and neuroanatomical structure of the brain, giving new voice to the intuitions so easily abandoned in the conventionalist interpretation of language in the first half of the last century. The limits of Babel, therefore, not only exist, but they can also be discovered in our flesh before any single experience: they are not the effect of an arbitrary convention – for a critical illustration, see [4] and references therein[1].

[1]The chosen technique to investigate the brain in the experiments described here is the so-called neuroimaging technique: in practice, the study of the metabolic activities of specific encephalic regions by measuring blood fluxes. The two main techniques are the Functional Magnetic Resonance (fMRI) and the position emission tomography (PET). It is important to beware of easy illusions. The research on neural networks with neuroimaging techniques can in some sense be compared to the attempt of reconstructing the map of the different cities of our planet having as unique data the flux of passengers at the airports: one can expect at the very best to have an approximate idea of the dimensions of the cities. The comparison is overly optimistic though: the number of possible circuits constituted by the hundreds of billions of neurons that on average

Finally, the third consequence consists on acknowledging that this linguistic model, based on the capacity to construct infinite structures starting from a finite set of elements, is unique of the human species. All living beings certainly communicate, but only human beings have this capacity for producing potentially infinite structures. Despite some surprising resistance, that this is the status of affairs is known at least since the seventies [10]. This conviction is such, for someone that studies the structure of communication codes, that it was the object of a plenary conference of the American Association of Linguistics [11], and even that, as it is easy to imagine, carries a definitively ecumenical character.

This characteristic of unicity, combined with the property of producing potentially infinite structures, has in turn a fundamental consequence that cannot be forgotten in any speculation about the evolution of language, or better about its filogenesis. Indeed, it must be clear that, being the capacity of producing potentially infinite structures, a specific character of human communication, it is theoretically inadmissible to affirm that there exists a gradual difference of this characteristic between the animal species: infinite, indeed, is either completely so, or not at all. One cannot just have a slice of infinity. Therefore, there cannot be similar languages to the human language, since any finite set, large as it may be, cannot be similar to infinity. Finally, another remark of the linguistic relativism, based not on rules but in the inventory of words: in the fifties of the last century a hypothesis took canonical form (which in ways more or less explicit was already circulating for quite some time), namely the idea that different languages correspond to different world visions due to the different vocabulary that each language possess (the so called Sapir-Work hypothesis).

## 4. Measuring the world vision

Let us emphasize: there are not only ways which are more or less effective in acting in the world — this is so obvious as stating that anyone trying to master a given technique must at the same time assume the basic wording — but there are also true and real different sensorial perceptions. It is not difficult to understand how, beyond this incarnation of relativism, was hidden the attempt, more or less explicit, of providing a gradation of merit between different languages, as if some of them were more suited to the perception of reality. Let me insist: perception. It is clear, for example, that in a language such as German, where building composite words is much more frequent than in Italian, one has more experience in building new suitable terms that allow to avoid periphrases and paraphrases, but from this to say that whoever speaks German sees a radio (or a dusk) in a different way as someone who instead speaks Italian is an unacceptable logical leap.

form an human brain is of the order of $10^{10^6}$: a network beyond imagination if one takes into account that the number of particles of which the universe is composed is around $10^{72}$ [9]. So it is little what can be seen, but it is nevertheless not negligible.

This implies that, on top of any ethical judgement, this hypothesis simply fails to reproduce the observed data. In the meantime, having a measure for the world vision is not possible: it does not exist even in theory a metric that allows us to understand if who speaks Italian or who speaks Tagalog perceive the world in different ways. It would be needed first of all to reach an agreement on what "world vision" really means. But in the very few cases where performing an experiment is acceptable, it shows that the world vision does not change by varying the language; if anything what might change is the interaction with the world. The example of the research of the colour names is paradigmatic in this respect. People called to distinguish different colours put over a table (without giving them a name) do not act in different ways: the perception remains the same even when the dictionary changes.

But let us mention that even this universalist vision has reductionist risks. We cannot forget that the scientific study of syntax is born in the second half of the last century to provide solely a description of the degree of variation of the class of human languages. The prediction of how and what an individual can say at a given moment, in a given context, outside banal cases, dares not enter in the research program nor be considered in a quantitative way, neither neuropsychologically nor at the molecular level: the linguistic creativity is not less true for this reason, but exactly as in the case of conscience, it is not measurable in quantitative terms.

This is not a nihilist resignation characteristic of weak thinking, though, just as it was not a nihilist resignation the decision of Newton to describe gravity as action at a distance, rejecting the contact mechanics of Descartes that was the orthodoxy at that time. It is an astonishing consequence that Chomsky refers precisely to Descartes when defining the fundamental capacity of human language — that of understanding and producing an infinite set of sentences — and that then he recognizes that in the very heart of language lays the mystery.

## References

1 Chomsky, N. (2004), The Generative Enterprise Revisited, (Mouton de Gruyter: Berlin – New York).
2 Greenberg, J.H., Ed. (1963), Universals of Language (The MIT Press: Cambridge, Mass.)
3 Rizzi, L. (2009), "The discovery of language invariance and variation, and its relevance for the cognitive sciences", Behavioral and Brain Sciences, 32, pp. 467-468.
4 Moro, A. (2008), The Boundaries of Babel. The Brain and the Enigma of Impossible Languages (The MIT Press: Cambridge, Mass.).
5 Bar-Hillel, Y. (1970), Aspects of Language (The Magnes Press: Jerusalem).
6 Chomsky, N. (1957), Syntactic Structures (Mouton: The Hague – Paris).
7 Chomsky, N. (1957), "Review of Skinner", Language, 35, pp. 26-58.

8 Lenneberg, E. (1967), Biological Foundations of Language (John Whiley & Sons: New York).
9 Edelman, G.M., G. Tononi (2000), A Universe of Consciousness: How Matter Becomes Imagination (Basic Books: New York).
10 Terrace, H.S., L.A. Petitto, R.J. Sanders & T.G. Bever (1979), "Can an Ape Create a Sentence?", Science, 206, pp.891-902.
11 Anderson, S.R. (2007), "The Logical Structure of Linguistic Theory", *president's speech at the 75th Annual Meeting of the Linguistic Society of America*, in Language, LXXXIV, 4, pp. 795-814.
12 Chomsky, N. (1966), Cartesian linguistics: a chapter in the history of rationalist thought (Harper & Row: New York).
13 Piattelli Palmarini, M. (2008), *Le scienze cognitive classiche: un panorama* (Einaudi: Torino).

euresis journal

euresis
Association for the Promotion of the Scientific Endeavour

Volume 3
Summer 2012

journal homepage
www.euresisjournal.org

# Is consciousness epiphenomenal? Social neuroscience and the case for interacting brains

John T. Cacioppo & Gary G. Berntson

Center for Cognitive and Social Neuroscience, 5848 S. University Avenue, University of Chicago, Chicago, IL 60637
Cacioppo@uchicago.edu

## Abstract

*Consciousness has been cast as epiphenomenal. We show that this conclusion is based on the analysis of consciousness in the context of the individual human brain. The premise of social neuroscience is that the study of the human brain is incomplete when treated as a solitary organ. When the brain is viewed within a social context — that is, within a context of interacting brains — the material mechanism by which consciousness plays a role in subsequent brain states is revealed.*

## 1. Introduction

Volition and consciousness have traditionally been important determinants in judgments of whether a person is culpable for an action. Volition, or free will, has been defined by philosophers as possessing the power to have done otherwise or to choose to do otherwise given the same circumstances. Free will, in turn, involves the notion that consciousness plays an instrumental role in what we feel, think, choose, and do. But what is the nature of consciousness? This is a question that has interested philosophers for centuries and one that scientists have now begun to ask. Rene Descartes' dualistic proposal that body and mind were distinct and separate had a profound effect on thought on the topic, including our notions of culpability. Dualism provided simple answers to these questions. Humans were cast as having a mechanical, determined, animalistic side and a conscious, volitional, spiritual side. Consciousness and free will were thought to be faculties of the latter, so the assignment of culpability required only that the offensive behavior not be a product of the reflexive, animalistic side of human nature.

Contrary to Descartes' approach, most scientific approaches to consciousness are grounded in monism and determinism. By monism, we mean that there is only one ultimate substance.

Monism does not mean that there is only one optimal *representation* of this substance, however. Different forms of representation make certain calculations easy and others difficult, as when using the periodic table to create new molecules versus using recipes to create dinner. In this context, monism means that the mind and consciousness are products of the operations of the central nervous system (CNS), which includes the brain and spinal cord but which for simplicity we will refer hereafter to as the brain.

By determinism we mean that there are specifiable conditions for everything that happens. Recent work in philosophy, psychology, psychiatry, and neuroscience has called into question the functional importance of consciousness and, with it, the notion of free will in human actions. The prediction of decisions based on brain functions observed prior to the decision has led to suggestions that the perception of free will is an illusion, that consciousness is epiphenomenal, and that people may not have as much responsibility for their actions as assumed. This notion is not entirely new, of course. As Hume observed, our actions are determined, in which case we are not responsible for them, or they are the result of random events, in which case we are not responsible for them.

The notion is that if consciousness is predetermined by antecedent brain functions, and if it can have no functional role in determining subsequent brain states, then consciousness would be without function –- that is, consciousness would be epiphenomenal. We will return to this thesis, but we begin by distinguishing among the concepts of free will, consciousness, mind, and brain. We then examine key elements of the somato-motor machinery through which human behavior is expressed. We then outline *within a monist, deterministic framework* how the brain and consciousness may have reciprocal influences. If true, then consciousness is not epiphenomenal but rather has an important functional role in the operation of mind and brain. We conclude by considering some of the implications of such a state of affairs, including the notion of free will.

## 2. Definitions

The constructs of free will, consciousness, mind, and brain are like a set of Russian dolls, with each construct fitting within the domain of the next. We may *will* to do this or that, which is to say we exercise choice. The key element of *free will*, however, is that we could choose to do otherwise given the same circumstances. One can have the conscious experience of will and choice, and the conscious conviction that one could have chosen otherwise, but if consciousness is epiphenomenal then the feeling of *free* will is an illusion.

Hume reasoned that if it were an illusion, one could not be morally responsible for one's actions. There are reasons to reject this conclusion, whether or not free will is an illusion, however. Society requires that people be responsible for their actions. Even the behavior

of nonhuman primates is influenced by the reactions of others to their own actions. If an individual acts in a way that violates the rules of the group, the individual may be held responsible for these actions, and the rewards that are lost or the costs that are incurred can reduce the likelihood of repeat offenses against the collective. Importantly, this is the case whether or not consciousness is epiphenomenal and whether or not free will is an illusion. The same behavior enacted by accident is treated differently because the offensive behavior is less likely to re-occur even in the absence of social costs. In this way, social responsibility for one's deliberate actions is a fundamental building block of society.

By consciousness, we mean awareness, including the sensations, perceptions, emotions, thoughts, and mental images on which a person can report. Among the set of events that constitute our consciousness is the experience that we freely choose among various ways to act. I may write an essay that advocates a position contrary to my beliefs because I was instructed to do so as part of an experiment in which I was participating -– in which case I would feel as if I had no choice, or because I chose freely to do so -– in which case I would feel as if I could have chosen otherwise. Research on cognitive dissonance theory has shown that people can be randomly assigned to "low choice" or "high choice" conditions, and all can be induced to write the same counterattitudinal essay. That is, the experimenter determines whether a participant writes the essay with the conscious experience of having no other option but to do so, or the participant writes the essay with the conscious experience of having freely chosen to do so. In neither case was their actual choice -– participants would not normally write such an essay without the experimenter's subtle manipulations of them. But whether or not they felt as if they had a conscious choice to do so greatly influenced what happened as a result of expressing the behavior. Participants in the low choice condition did not change their beliefs or behaviors toward the position they advocated in the essay because they felt no responsibility for this behavior. Participants in the high choice condition, however, changed their beliefs and attitudes to align them with the position they advocated in the essay because they consciously felt responsible for this behavior. This result has been interpreted to mean that conscious choice matters, even when it is an illusion (see review by [1]). But consciousness itself may be epiphenomenal, in which case conscious choice — or the (perceived) expression of free will — does not really matter, either – it would be the underlying brain states that matter.

Consciousness, then, includes the experience of free will but is not limited to free will. The mind, in contrast, refers to the structures and processes responsible for thought, emotion, and behavior. As such, the mind encompasses but extends beyond consciousness to include unconscious processes. For instance, Figure 1 illustrates the Stroop Test, a task that illustrates the influence of nonconscious processes on our thoughts and behaviors. The instruction to participants is simple: Identify the color of ink in which each letter string is printed. Words are then presented, one at a time, and participants announce the color in which the word is presented as quickly as they can. When the words are the names of colors (e.g., blue, red, or-

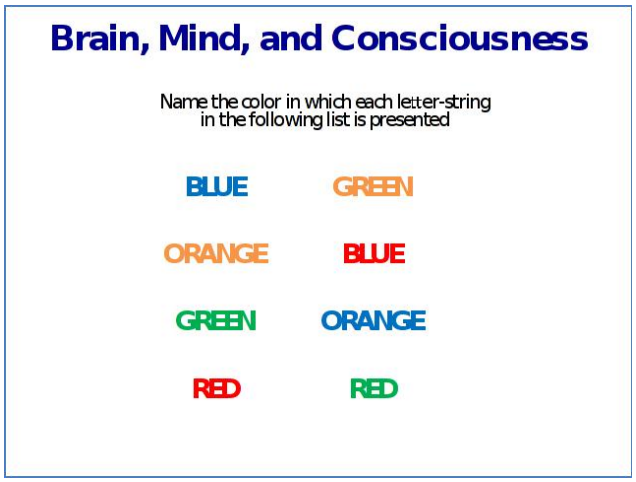


Figure 1: *The Stroop Test.*

ange) and are printed in the corresponding color (e.g., "blue" is printed in blue), participants perform the task quickly and with few if any errors (see Figure 1, left column). However, when the words are the names of colors printed in discrepant colors (e.g., "blue" is printed in red), participants stumble, taking longer to identify the colors and making more errors in identifying the colors in which the words are printed (see Figure 1, right column). The difference in behavior is because the brain is reading the words even though the participants are not consciously choosing to do so and, in fact, are not even consciously aware of doing so. When the words and colors are consistent, nonconsciously attending to the meaning of the words facilitates performance on the task. When the words and colors are inconsistent, however, this nonconscious information processing interferes with the conscious attempt to identify the color in which the words are printed. There are many such examples that could be given for preattentive or nonconscious processes of the mind, but the Stroop test is sufficient to illustrate that the mind includes conscious experience but its operations are not limited to conscious experiences [2].

Finally, the brain is the organ of the mind but also performs functions that are not directly relevant, including autonomic, neuroendocrine, and immune function such as the modulation of the regulatory mechanisms involved in the maintenance of homeostatic set-points for body temperature, blood pressure, and blood sugar. If the brain were to fail in these operations, there would be dramatic consequences for the mind, consciousness, and behavior but the same could be said about the failure of any major organ (e.g., the heart). Thus, free will (whether an illusion or not) represents an experience that falls within the larger domain of consciousness, consciousness (whether epiphenomenal or not) represents a set of structures and processes within the larger domain of the mind, which in turn represents a set of structures and processes within the larger domain of the brain.

## 3. The machine

A common misconception about human behavior is that it follows in a straightforward fashion from beliefs and intentions, enacted by the homunculus positioned in the primary motor cortex. Dennett 1991 [3] describes such an enactment as the Cartesian theater and regards it as a remnant of instinctively dualistic thinking. One problem with this conception is that it accounts for human agency and behavior in a circular fashion [4]. A second is that the primary motor cortex is only a small piece of a heterarchically organized motor system. A third is that the cell bodies of the motor system themselves represent but a small part of the complex information processing capacities of the brain (and CNS more generally).

Behavior is complexly determined. Neurobiologically, human action is enacted through the motor system, which controls somatic muscles of skeletal movement. The muscles are innervated by the lower and upper motoneuron systems. The lower motoneuron system, sometimes called the final common pathway, consists of cell bodies located in the spinal cord, brain stem, and pons with axon fibrils extending to motor end plates on muscle fibers. A lesion of any part of the lower motorneuron system results in a flaccid paralysis (loss of motor control and muscle tone) of the muscles to which the severed motoneurons had projected.

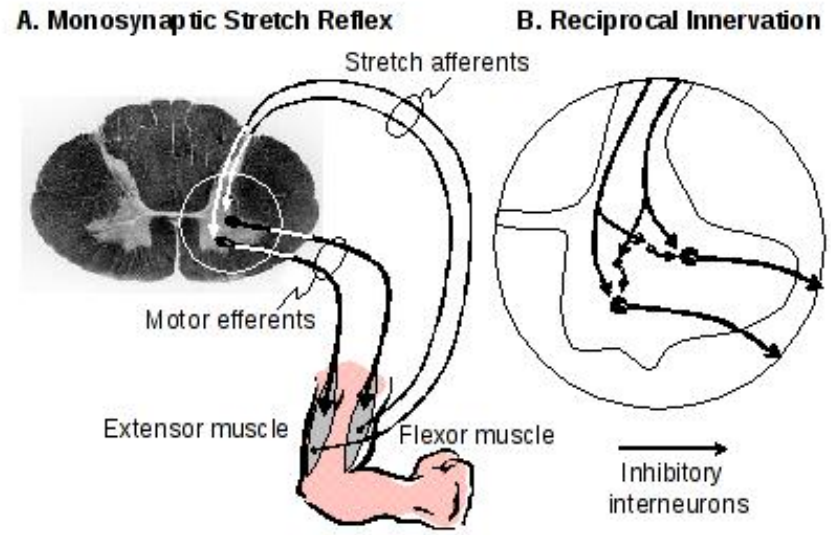


Figure 2: *Stretch Reflex. A. Basic parallel circuits of the flexor and extensor stretch reflex. B. Reciprocal innervation, an example of Sherrington's alliance of reflexes. Dotted lines represent inhibitory interneurons which achieve a level of reciprocal integration between flexor/extensor motor neuron pools. (From [7].).*

As documented by early investigators, such as [6], basic somatomotor control is effected at the level of the spinal cord, with spinal reflexes representing the lowest central level in

somatomotor control systems (Figure 2). Spinal reflex circuits are relatively simple and may be comprised of a single central synapse. The monosynaptic stretch reflex is exemplified by the knee jerk reflex to the physician's rubber mallet. This reflex entails an afferent somatosensory link arising from muscle stretch receptors, which synapse directly on the lower motor neurons controlling that muscle. This simple circuit provides for a reflexive contraction of the stretched muscle, which tends to compensate for the perturbing stretch. Stretch reflexes exist in all major classes of somatic muscles, including flexor and extensor muscles.

Opponent flexor (e.g., biceps) and extensor (e.g. triceps) reflexes antagonize one another and promote opposite outcomes for the limb (flexion and extension, respectively). The basic neural circuits of these reflexes are independent and organized in parallel, and they have limited inputs and outputs, allowing for rapid, efficient processing (Figure 2). The cost of this efficiency, however, is that lower-level systems have limited integrative capacity. Moreover, they can be in conflict. Simultaneous stretch of both the flexor and extensor muscle may lead to a reflexive increase in muscle tension in both muscles, but because they are opposed in their actions, there may be no resultant limb movement.

Greater levels of integration in motor systems are achieved by hierarchical circuits that promote coordination among the basic spinal reflexes – what Sherrington referred to as the *alliance of reflexes*. In our flexor/extensor example, this entails a collateral projection of the stretch receptor afferents onto inhibitory interneuron circuit elements, which in turn project to and inhibit the motor neuron for the opposing muscle. Stretching the flexor muscle, for example, results not only in activation of the flexor motor neurons (stretch reflex), but also inhibition of the opposing motor neurons via an inhibitory neural pathway. This exemplifies a general principle of neural organization articulated by Sherrington — the principle of *reciprocal innervation* — which stipulates that neural systems promote specific outcomes by activating the mechanisms for the target response while at the same time inhibiting opposing responses [7, 8].

Sherrington's alliance of reflexes does not stop with reciprocal innervation. The cell bodies of the lower motoneuron system are innervated by upper motoneurons from the pyramidal and the extrapyramidal systems. A lesion of each of these upper motoneuron systems has distinctive behavioral effects. The cell bodies of the pyramidal upper motoneuron system are located in the precentral gyrus of the frontal lobe, and long axons project to the cell bodies of the lower motoneuron system in the pons, brain stem, and spinal cord. The pyramidal motor system controls all our voluntary movements –- so called because movements enacted through the pyramidal system tend to be associated with the conscious experience of having chosen to make the movement, even when this conscious experience trails measurable brain activities predicting the movement and precursors of the movement itself by seconds (e.g., [9]). For instance, a lesion of the pyramidal tract from the facial region of the motor cortex to the cell bodies in the facial nerve nucleus results in the inability to volitionally smile on

the contralateral side of the face. If the afflicted individual hears a humorous joke, however, a normal smile is produced [34].

The ability exists to produce a symmetrical smile spontaneously/reflexively but not volitionally because the extrapyramidal upper motoneuron system is still intact. The extrapyramidal upper motoneuron system is phylogenetically older than the pyramidal system and thus plays a relatively more important role in lower animals. The nuclei constituting the extrapyramidal system are manifold with the major parts located in the basal ganglia, including the caudate, putamen, and globus pallidus; the thalamus and subthalamic nucleus; and the substantia nigra and red nucleus in the midbrain. All of these nuclei are connected synaptically to one another, the brain stem, cerebellum, and the pyramidal system. A lesion of the extrapyramidal innervation of the facial nerve nucleus leaves the afflicted individual capable of producing a voluntary symmetrical smile, but the individual is incapable of producing a normal spontaneous smile when they hear something humorous [34].

As hierarchical levels are layered on the motor control system, progressively higher levels receive a wider array of inputs, have greater circuit complexity and computation capacity, and can achieve a broader and more flexible range outputs [7, 8]. At the highest levels, beyond the primary motor cortex, cerebral systems must process a tremendous amount of sensory information, and integrate this information with associative networks, emotional/motivational substrates, and expectancies, in the contexts of strategic goals and tactical plans. This requirement for enhanced information processing can impose a processing bottleneck that necessitates a slower, more serial mode of processing and selective attentional mechanisms, e.g. [10]. Although it is these highest level systems that confer the greatest cognitive and behavioral capacity, they do not operate in isolation but depend upon and interact with lower levels in the hierarchy.

Hierarchical dimensions of central nervous system organization can be demonstrated anatomically as well as functionally –- see [5]. The simple hierarchy depicted by the solid lines in Figure 3, however, belies the true complexity of neurobehavioral substrates, as long ascending and descending pathways (dashed lines) can bypass intermediate levels of hierarchical organization and interconnect across widely separated neural levels. Cortical motor neurons project not only to intermediate-level somatomotor networks, but also directly onto spinal motor neurons through long descending pathways [11, 18]. The long ascending and descending pathways in neural hierarchies, together with the existence of lateral interactions among elements (such as those that underlie reciprocal innervation), yield what has been termed a heterarchical organization -– see [5]. The outputs of a strict hierarchical system are coherent, as all levels are linked by intermediate regulatory levels, and all outputs are by final common pathways. In a heterarchical organization, however, higher levels can directly access output mechanisms independent of intermediate levels. This organizational feature allows for concurrent expression of multiple re-representative systems, which can increase

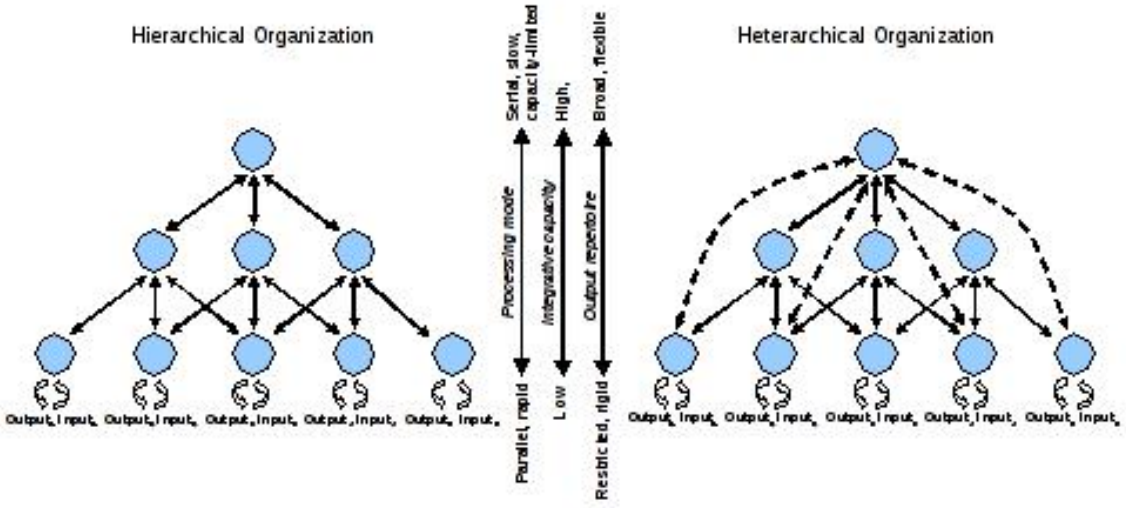


Figure 3: *Hierarchical and heterarchical organizations. A heterarchy differs from a hierarchy in the existence of long ascending and descending pathways that span intermediate levels. Properties of the levels in both classes of organizations lie along the illustrated continuua of processing mode, integrative capacity and output repertoire. Heterarchical organizations, however, have greater integrative capacity and output flexibility as the long ascending and descending projections provide inputs and outputs that are not contrained by intermediate levels. (From [7].).*

behavioral complexity but can also lead to functional conflicts. By volitionally stiffening the leg, for example, higher level systems can mask the stretch reflex and usurp control of motor neuron pools.

Paralleling the pattern for somatic motor control, adaptive and protective reflexes and mechanisms are organized at all levels of the neuraxis. The pain withdrawal reflex, for example, is organized at the level of the spinal cord and can be seen even after spinal transections that isolate the cord from higher brain systems. Pain withdrawal reflexes are protective reactions that arise from somatosensory afferents, carrying nociceptive signals, which anatomically link to flexor neuron pools by a multisynaptic spinal pathway. Through this spinal reflex circuit, noxious stimuli yield a protective flexor withdrawal response. Likely because of their adaptive value, pain withdrawal reflexes are among the earliest to develop and the most resistant to disruption. Although pain withdrawal reflexes at the level of the cord may not require the invocation of a construct of emotion or affect, they represent an important low level evaluative mechanism for escape from noxious stimuli. Moreover, despite their neural simplicity, these circuits can show operant conditioning of escape which can support not only escape from, but active avoidance of, pain stimuli [4].

In contrast to the primitive avoidance system associated with flexor reflexes, separate mechanisms exist at the level of the cord for opposing extensor responses which promote engagement with the environmental stimuli. Reflexes such as the extensor reflex in response to non-painful cutaneous stimulation of the palm or the sole of the foot contribute to postural,

locomotory and grasping responses that serve to engage the organism with the environment.

As considered above, flexor and extensor reflexes are organized largely in parallel, as they control distinct motoneuron pools for opposing muscles. Nevertheless, they do interact. These lower reflex substrates are integrated by higher-level circuits, such as those that implement reciprocal innervation, which tends to reduce concurrent activation. They are also impacted by even higher systems that contribute to volitional actions and confer a greater degree of flexibility and control over flexor and extensor motor neuron pools and lower reflex substrates. Thus, we can volitionally contract both flexor and extensor muscles (e.g., in stiffening the arm) which can overcome the lower level reciprocal innervation, and we are able to override or suppress flexor pain withdrawal reflexes (e.g., to remove a sliver from the finger).

Given there are multiple levels of neurobehavioral control within the CNS and the responses these nuclei activate can differ in a given circumstance (e.g., as when the point of a needle containing a life-saving vaccine punctures the skin), response conflicts can occur. The frontal regions have long been thought to be involved in executive functions such as dealing with response conflicts, formulating goals and plans, selecting among options to achieve these goals, monitoring the consequences of our actions in light of our goals, and inhibiting, switching and regulating our behaviors accordingly. Aron, 2008 [12] reviews evidence that the initiation of a motor response proceeds from the planning areas of the frontal cortex to the putamen, globus pallidus, thalamus, primary motor cortex, motor nucleus in the spinal cord, and finally to the muscles. Being able to inhibit a motor response once it has been initiated has obvious adaptive value, and this inhibition involves the right inferior frontal cortex, which projects to the subthalamic nucleus (a region of the basal ganglia that may act on the globus pallidus to block the motor response). Monitoring for response conflicts, in turn, appears to involve the dorsal anterior cingulate and the adjacent presupplementary motor area which, in turn, is connected to the right inferior frontal cortex and subthalamic nucleus. Switching also involves the presupplementary motor area and right inferior frontal cortex. Together, this work has led to a model in which the presupplementary motor area may monitor for conflict between an intended response and a countervailing signal and when such conflict is detected the "brakes" could be put on via the connection between right inferior frontal cortex and the subthalamic nucleus region [12]. The term "monitoring" can be misleading, however. There is no homunculus overseeing our actions, but rather in a quite deterministic fashion the activation of multiple, sometimes conflicting, response circuits that include excitatory and inhibitory connections leads to the activation of some responses over others based on which first reaches its response threshold.

There is also an extensive behavioral literature showing that behavioral intentions predict volitional behaviors, and that the experimental manipulation of these behavioral intentions leads to corresponding changes in volitional behavior. In the theory of planned behavior

[13], intentions are determined by three distinct constructs: attitude toward the behavior, subjective norms about the behavior, and perceived behavioral control. Attitude, in turn, is determined by *beliefs* ($b_i$) and *evaluations* ($e_i$) of those beliefs such that the *attitude* = $\sum_i b_i e_i$ for $i = 1, n$. Subjective norm is determined by *normative beliefs* ($nb_j$) and the *motivation to comply* with these beliefs ($mc_j$) such that the *subjective norm* = $\sum_j nb_j mc_j$ for $j = 1, m$. And finally, the perceived behavioral control is determined by *control beliefs* ($cb_k$) and the *perceived probability* that these beliefs are true ($pp_k$) such that *perceived behavioral control* = $\sum_k (cb_k pp_k)$ for $k = 1, o$.

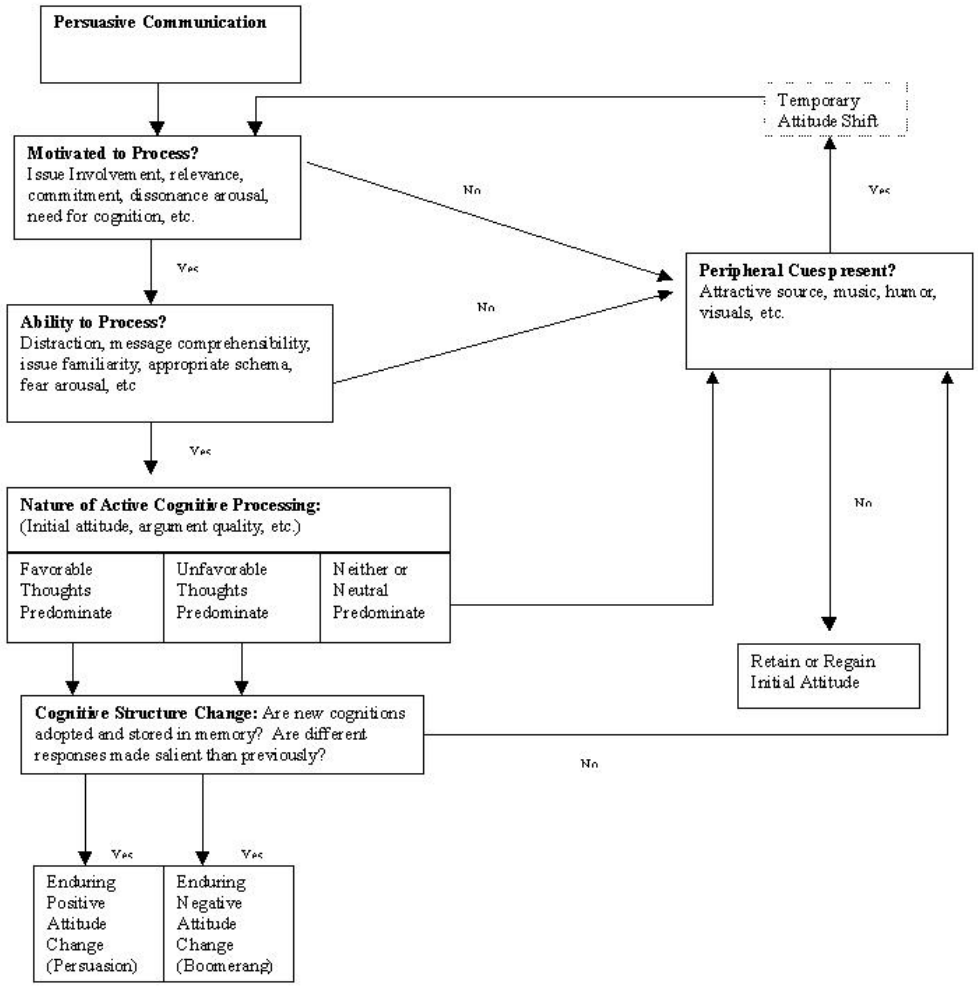


Figure 4: *The Elaboration Likelihood Model. (From [14].).*

Where do the beliefs and evaluations of beliefs that lead to the attitude come from? Hovland and colleagues suggested that attitudes were the result of comprehending and encoding the message arguments (i.e., belief statements). This view proved to be incorrect, and the message learning approach was replaced by the elaboration likelihood model (ELM) of attitude

change ([1, 14]; Figure 4). The ELM specifies that there are both effortful (central route) and non-effortful (peripheral route) attitude change processes, that attitude change derived through the central route can reflect rational or psycho-rational issue-relevant thinking, and that the central and peripheral route to attitude change operate under different but specifiable conditions. Finally, according to the ELM, there are different consequences of thoughtful and non-thoughtful persuasion such as attitudes derived through the central route are more enduring, more resistant to counterpersuasion, and more predictive of behavior than are attitudes derived through the peripheral route. There is now an extensive behavioral literature supporting the predictions of the ELM. For instance, multiple arguments can produce thoughtful change under some conditions (e.g., under high personal relevance of the message people scrutinized the merits of the arguments), but can produce the same amount of unthoughtful change under other conditions (e.g., under low personal relevance of the message people based their attitudes simply on the number of arguments presented regardless of their merit). The former thoughtful attitude change, however, is more persistent and predictive of behavior than the latter non-thoughtful change.

In the past decade, theory and research has led to an appreciation for the importance of the role of automatic associative processes, or implicit attitudes, especially in the production of spontaneous behaviors, e.g. [30, 16]. These have been termed "implicit" because they do not involve consciousness, at least in the same way as beliefs, attitudes, intentions, and decisions. Implicit processes have traditionally been depicted as representing a single, unitary mechanism, but the heterarchical organization outlined here suggests these processes, though no less deterministic, are multifarious in underlying mechanism and function.

## 4. The conundrum

The work on implicit attitudes and behaviors expands rather than contradicts work on the experience of agency (i.e., the awareness of being in control both of one's own actions and through them of events in the external world). The prediction of decisions based on brain functions observed prior to the decision have led to suggestions that people may not have as much responsibility for their actions as assumed, however. The argument for consciousness being epiphenomenal can be parsed into two parts. The first part is that consciousness is fully predicted and determined by the brain's prior and ongoing states. The second part is equally important, however, and it is the notion that consciousness plays no role in subsequent brain states. From the perspective of the brain, it is difficult to conceive through what material mechanism consciousness could conceivably alter subsequent brain states above and beyond the prior and ongoing brain states that give rise to consciousness. This is what philosophers call overcausation. If there is a sufficient causal explanation in the physical domain, there cannot be a meaningful explanation in some other domain. Without the specification of such a mechanism, one is left to conclude that consciousness has no functional role and,

hence, is epiphenomenal.

The argument is outlined more formally in Equation 4.1. The quantity $\beta_t$ represents the brain state at any given point in time, designated as $t$. This brain state is designated to be a function of a constant representing the unique capacities of a given individual's brain, the observed brain state at time $t-1$, and the prediction error by the brain at time $t-1$, which represents the difference between the state of the world (and brain) predicted for time $t-1$ and the state of the world (and brain) at time $t-1$. Our brains are metabolically expensive. They represent approximately 5% of our body mass but are responsible for approximately 20% of our oxygen consumption. This prediction error can be viewed as operating like a compression algorithm to minimize these metabolic costs. To send a video signal digitally, for instance, the information in each pixel of each frame of a film can be transmitted, but this is resource intensive and expensive. To make this transmission more efficient, one can send each pixel of the first frame and the pixels that have changed in each succeeding frame. Using such a compression algorithm greatly reduces the transmission load and costs. Of course, to know which pixels have changed from frame to frame, one has to compare what is predicted at each pixel based on the prior frame and what is observed. If there is a difference, then the new state of the pixel is transmitted and the image is updated. The use of such compression algorithms provides significant savings in terms of processing load to achieve transmission of the same information. Dating back to the classic work of Rescorla and Wagner (1968) [33], learning theorists have emphasized prediction error as an important component of updating our cognitive representations of the world and improving our response repertoires.

$$\begin{aligned} \text{Given } \beta_t &= f(\mu + \alpha_{i,t-1}\beta_{t-1}^{obs} + \alpha_{j,t-1}(\beta_{t-1}^{obs} - \beta_{t-1}^{pred}) + ...), \\ \text{Determinism: } &\beta_t \rightarrow \beta_{t+1} \rightarrow \beta_{t+2} \rightarrow \beta_{t+3}... \\ \text{Monism: } &\overset{\downarrow}{\chi_t} \quad \overset{\downarrow}{\chi_{t+1}} \quad \overset{\downarrow}{\chi_{t+2}} \quad \overset{\downarrow}{\chi_{t+3}} \ ... \end{aligned} \tag{4.1}$$

- Cases to consider:
  1. walking along a smooth sidewalk;
  2. walking along a smooth sidewalk when you surprisingly step into an unforeseen depression, leading you to verbally express an expletive ($\chi_{t+1}$).
- Importantly, in no instance does the portion of consciousness, $\chi$, even when acute and expressed as an expletive have any influence, directly or indirectly, on subsequent brain states, leading to the conclusion that consciousness is epiphenomenal.

As also depicted in Equation 4.1, brain state at time $t$ predicts, and in fact determines, the brain state at time $t+1$, and so on. This follows from determinism, a basic neuroscientific tenet. Determinism has been challenged, for instance, based on arguments that quantum mechanics have shown events are not deterministic and, therefore, brain states cannot be

fully determined by prior brain states. Pinker (2003) [18] has noted that the neural operations of the brain are at a sufficiently macroscopic level that the uncertainty of quantum physics does not translate into the same uncertainty at the level of neuronal activation. With approximately 100 billion neurons and an estimated 1014 neuronal connections in the brain, we are far from being able to specify the precise mechanism through which each brain state determines the succeeding. However, scientific determinism does not mean one has a complete explanation, only that such an explanation exists.

The bottom line in Equation 4.1 illustrates how consciousness at time $t$, depicted as $\chi_t$, is a function of the brain state at time $t$. Consciousness, in short, is a product of the functions of the brain –- consistent with the principle of monism. As such, however, it is left with no more function than the output display on a computer screen, playing no role in determining the succeeding state of the computer that produced this output.

Consider the case of an individual walking along a straight, smooth path. Each step is orchestrated by the brain and is a close replica of the preceding step. There is nominal prediction error, so $\beta_t$ determines $\beta_{t+1}$ which determines $\beta_{t+2}$ and so forth. The corresponding state of consciousness at each point in time can be expressed as $\chi_t$, $\chi_{t+1}$, $\chi_{t+2}$, and so on, but there is no known mechanism through which consciousness can influence the succeeding brain state. Although counterintuitive and perhaps implicitly objectionable, this line of reasoning leads to the conclusion that consciousness is epiphenomenal.

Now consider the case of an individual walking along a straight, smooth path who is surprised when a step falls into unnoticed depression along the path, leading the person at time $t+1$ to express of what they suddenly became aware, "now that was a surprise," represented by $\chi_{t+1}$. The difference between this case and the prior case is that there is a significant prediction error at time $t+1$, and with this prediction error comes a heightened sense of awareness of some of the events that the brain was processing. Despite these differences in brain states and corresponding conscious states, the prediction error is included in the definition of $\beta_t$ so, as in the prior case, $\beta_t$ determines $\beta_{t+1}$ which determines $\beta_{t+2}$ and so forth, and the corresponding state of consciousness at each point in time, expressed as $\chi_t$, $\chi_{t+1}$, $\chi_{t+2}$, and so on, is determined by the corresponding brain state. In this case, $\chi_{t+1}$ is associated with the statement, "now that was a surprise," but this difference is of no real consequence. There still is no known mechanism through which consciousness or this expression could influence the succeeding brain state. Again, then, this line of reasoning leads to the conclusion that consciousness serves no function.

## 5. The solitary brain versus the social brain

The analyses of consciousness to date have treated the solitary human brain as the appropriate unit of analysis. However, we are a social species, and social species, by definition, form organizations that extend beyond the individual. These structures evolved hand in hand with behavioral, neural, hormonal, cellular, and genetic mechanisms to support them because the consequent social behaviors helped these organisms survive, reproduce, and care for offspring sufficiently long that they too reproduced, thereby ensuring their genetic legacy. Social neuroscience is the interdisciplinary academic field devoted to understanding how biological systems implement social processes and behavior, and how these social structures and processes impact the brain and biology. Social neuroscience is not a cognitive neuroscience of social stimuli.

Behavioral neuroscience is a perspective in which the nervous system and brain are viewed as instruments of sensation and response. Research representing this perspective tends to focus on topics such as learning, memory, motivation, homeostasis, sleep and biological rhythms, and reproduction –- and on the neural mechanisms underlying these behavioral functions. Cognitive neuroscience emerged as a distinct functional perspective in which the brain is viewed as an information processing organ, with a focus on topics such as attention, perception, representations, decision-making, memory systems, heuristics, reasoning, and executive functioning – and on the neural mechanisms in the human brain that underlie these representations and processes [19]. Social neuroscience represents yet another broad perspective that extends beyond the structure and function of a single organism to investigate the functions that are altered by or are derived from the association or interaction of conspecifics (imagined or real)-– and on the neural and hormonal mechanisms underlying these structures and functions [20, 26]. If cognitive neuroscience is equivalent to the study of a computer connected to an electrical outlet, social neuroscience is equivalent to the study of a mobile, broadband connected computer linked to countless others via the internet.

Human social processes were once thought to have been incidental to learning and cognition, whereas the social complexities and demands of primate species are now thought to have contributed to the evolution of the neocortex and various aspects of human cognition. According to Dunbar and colleagues, e.g. [21], deducing better ways to find food, avoid perils, and navigate territories has adaptive value for large mammals, but the complexities of these ecological demands are no match for the complexities of social living (especially in hostile between-group social environments), which include: recognizing ingroup and outgroup members; learning by social observation; recognizing the shifting status of friends and foes; anticipating and coordinating efforts between two or more individuals; using language to communicate, reason, teach, and deceive others; orchestrating relationships, ranging from pair bonds and families to friends, bands, and coalitions; navigating complex social hierarchies, social norms and cultural developments; subjugating self-interests to the interests of

the pair bond or social group in exchange for the possibility of long term benefits for oneself or one's group; recruiting support to sanction individuals who violate group norms; and doing all this across time frames that stretch from the distant past to multiple possible futures. Consistent with this reasoning, human toddlers and chimpanzees have similar cognitive skills for engaging the physical world but toddlers have more sophisticated cognitive skills than chimpanzees for engaging the social world [4]; cross-species comparisons have revealed that the evolution of large and metabolically expensive brains is more closely associated with social than ecological complexity [23]; and a composite index of sociality in troops of baboons has been found to be highly correlated with infant survival [6].

Our survival depends on our connection with others. Born to the most prolonged period of utter dependency of any animal, human infants must instantly engage their parents in protective behavior, and the parents must care enough about their offspring to nurture and protect them. If infants do not elicit nurturance and protection from caregivers, or if caregivers are not motivated to provide such care over an extended period of time, then the infants will perish along with the genetic legacy of the parents [24]. Our developmental dependency mirrors our evolutionary heritage. Hunter/gatherers did not have the benefit of natural weaponry, armor, strength, flight, stealth, or speed relative to many other species. Human survival depended on *collective* abilities, not on individual might. C*ommunication* is critical to organizing these collective abilities.

Which brings us to two additional cases to consider in our investigation of the function of consciousness. First, consider an individual walking along a straight, smooth path who *imagines* another individual surprisingly stepping out from behind an obstruction, leading the first person at time $t + 1$ to express of what they suddenly became aware, "now that was a surprise," represented by $\chi_{t+1}$ (see Equation 4.2). As in the prior case, there is a significant prediction error at time $t + 1$, and with this prediction error comes a heightened sense of awareness of some of the events that the brain was processing. Despite these differences in brain states and corresponding conscious states, the prediction error is included in the definition of $\beta_t$ so, as in the prior case, $\beta_t$ determines $\beta_{t+1}$ which determines $\beta_{t+2}$ and so forth, and the corresponding state of consciousness at each point in time, expressed as $\chi_t$, $\chi_{t+1}$, $\chi_{t+2}$, and so on, is determined by the corresponding brain state. The other individual's brain states and consciousness are also constructs of the first person's brain states — so designated by their enclosure in the thought-bubble — so in fact these do not represent the true brain or conscious states of another individual, but only the constructions of a solitary brain. Accordingly, $\chi_{t+1}$ is associated with the statement, "now that was a surprise," but the expression again has no real consequence, and we are left to conclude that consciousness serves no function.

$$\begin{aligned}
\text{Given } \beta_t &= f(\mu + \alpha_{i,t-1}\beta_{t-1}^{obs} + \alpha_{j,t-1}(\beta_{t-1}^{obs} - \beta_{t-1}^{pred}) + ...), \\
\text{Determinism: } & \beta_t \rightarrow \beta_{t+1} \rightarrow \beta_{t+2} \rightarrow \beta_{t+3} ... \\
\text{Monism: } & \overset{\downarrow}{\chi_t} \quad \overset{\downarrow}{\chi_{t+1}} \quad \overset{\downarrow}{\chi_{t+2}} \quad \overset{\downarrow}{\chi_{t+3}} \ ... \\
\text{Imaginary Process: } & \begin{cases} \chi_t \quad \chi_{t+1} \quad \chi_{t+2} \quad \chi_{t+3} \cdots \\ \overset{\uparrow}{\beta_t} \rightarrow \overset{\uparrow}{\beta_{t+1}} \rightarrow \overset{\uparrow}{\beta_{t+2}} \rightarrow \overset{\uparrow}{\beta_{t+3}} \ ... \end{cases}
\end{aligned} \tag{4.2}$$

- Cases to consider:
  1. walking along a smooth sidewalk when *you imagine* an unforeseen person surprisingly step out from behind an obstruction, leading you to say "that was a surprise" ($\chi_{t+1}$)

Finally, consider the case of an individual walking along a straight, smooth path when another individual surprisingly steps out from behind an obstruction, leading the first person at time $t+1$ to express of what they suddenly became aware, "now that was a surprise," represented by $\chi_{t+1}$. The brain and conscious states of the second person (shown in Equation 4.3) follow the same principles as those of the first person, adhering to determinism and monism. As in the prior case, there is a significant prediction error at time $t+1$ when the first person encounters the second, and with this prediction error comes a heightened sense of awareness of some of the events that the brain was processing. Unlike the prior cases, however, the verbal expression of the part or whole of the contents of $\chi_{t+1}$, as "now that was a surprise" serves as a communication, thereby producing an effect on the succeeding brain state of the second person (depicted in Equation 4.3). The brain of the first person continues to predict the likely succeeding brain state, including in this case an attempt to predict the thoughts, feelings, and behavior of the second individual -- efforts that have been referred to as "mentalizing" and "theory of mind." Typically, the second person's conscious and behavioral responses are not perfectly predictable, however. In this particular case, the second person may respond to the communication, "now that was a surprise" with an apology for surprising the individual, an insulting expletive, a request for directions, and so forth. This person's response to the expression of the first individual's conscious content therefore influences that individual's subsequent brain state (depicted in Equation 4.3). In this instance, the portion of consciousness, $x_{t+1}$, expressed as surprise, has an indirect but deterministic influence on that individual's subsequent brain states. This line of reasoning implies that consciousness can play a deterministic, monistic, and functional role in brain states.

$$\text{Given } \beta_t = f(\mu + \alpha_{i,t-1}\beta^{obs}_{t-1} + \alpha_{j,t-1}(\beta^{obs}_{t-1} - \beta^{pred}_{t-1}) + ...),$$

$$\text{Determinism:} \quad \beta_t \rightarrow \beta_{t+1} \xrightarrow{\chi'_{t+1}} \beta_{t+2} \rightarrow \beta_{t+3}...$$

$$\text{Monism:} \quad \overset{\downarrow}{\chi_t} \quad \overset{\downarrow}{\chi_{t+1}} \quad \overset{\downarrow}{\chi_{t+2}} \quad \overset{\downarrow}{\chi_{t+3}} \ ... \tag{4.3}$$

$$\text{Second individual:} \begin{cases} \chi'_t & \chi'_{t+1} & \chi'_{t+2} & \chi'_{t+3}\cdots \\ \overset{\uparrow}{\beta'_t} \xrightarrow{\chi_{t+1}} & \overset{\uparrow}{\beta'_{t+1}} \rightarrow & \overset{\uparrow}{\beta'_{t+2}} \rightarrow & \overset{\uparrow}{\beta'_{t+3}} \ ... \end{cases}$$

- Cases to consider:
  1. walking along a smooth sidewalk when an unforeseen person surprisingly steps out from behind an obstruction, leading you to say "that was a surprise" ($\chi_{t+1}$), to which the person responds (unpredictably) by apologizing by surprising you, expresses an expletive, asking you directions, walking by without expressing anything, etc.
- Importantly, in this instance the portion of consciousness ($\chi_{t+1}$), expressed as surprise, has an indirect influence on subsequent brain states, meaning that conciousness has a monistic, deterministic, and functional role in brain states.

Consciousness, in short, is a product of the functions of the brain. When focused on the solitary brain, consciousness is left with no more function than the output display on a computer screen, playing no role in determining the succeeding state of the computer that produced this output. Of course the display appears on a computer screen because such computers are not designed to operate in isolation but rather are designed to make it possible for humans to interact with them. When a human operator is sitting at the keyboard, the computer display plays a crucial role in determining what are the subsequent states of the computing device as it influences what the human operator instructs the computer to do next. Eliminate the computer display, and the subsequent states of the computer are quite different than they would be if the human operator were able to see the computer display. We have posited an analogous argument for the functional yet deterministic and monistic role of consciousness – an argument that follows from the observation that we are a fundamentally social species, which is to say that our brains have evolved to interact with other brains.

## 6. Conclusion

The premise of social neuroscience is that the study of the human brain is incomplete when treated as a solitary organ. When the brain is viewed within a social context — that is, within a context of interacting brains — the material mechanism by which consciousness plays a role in subsequent brain states is revealed. From the perspective of social neuroscience, language evolved to communicate and coordinate with conspecifics rather than simply to talk to oneself. When communicating with others, our brain spontaneously attempts to predict

the present and future behavior of others through processes following under terms such as social cognition, mentalizing, and theory of mind. Although perhaps better than chance, we are far from perfect mind-readers. This imperfection means that when we communicate our conscious states to others, our subsequent brain states are not entirely predicted by the brain states that gave rise to our conscious expression. That is, our subsequent brain states are in part determined by our prior brain states and in part by the brain states that result from interactions with other brains – other brains whose influence on our subsequent brain states are themselves determined in part by the communication of aspects of our consciousness. One might counter that a person's communication is itself the result of prior and ongoing brain states so that it is the brain states, and not aspects of consciousness, that are being communicated. The case of the display screen on a computer helps clarify why this output is instrumental even if its influence is mediated through interactions with other humans. Thus, because our brain underlies communication with other brains – that is, because it is social, the conscious beliefs and intentions we communicate to others has an impact on the brains of others that were not entirely predictable by their prior brain states, and their responses to us influence our subsequent brain states in ways not entirely predictable by our prior brain states. Consciousness in this social context may therefore have the potential to serve a functional role.

## 7. Acknowledgements

This research was supported by National Institute of Aging Grant No. RO1 AG034052-01.

## References

1 Petty, R. E., and J.T. Cacioppo (1981), Attitudes and persuasion-classic and contemporary approaches (W. C. Brown Co. Publishers: Dubuque, IA).
2 Öhman, A., A. Flykt, and D. Lundqvist (2000), "Unconscious emotions: Evolutionary perspectives, psychophysiological data, and neuropsychological mechanisms", in: Cognitive neuroscience of emotion, Eds. R. Lane & L. Nadel (O.U.P.: New York), pp. 296-327.
3 Dennett, D. C. (1991), Consciousness explained (Little, Brown & Co. USA).
4 Grau, J., E. Crown, A. Ferguson, S. Washburn, M. Hook, and R. Miranda (2006), "Instrumental learning within the spinal cord: underlying mechanisms and implications for recovery after injury", Behavioral and cognitive neuroscience reviews, 5(4), 191-239.
5 Berntson, G. G., J.T. Cacioppo, and K.S. Quigley (1993), "Cardiac psychophysiology and autonomic space in humans: empirical perspectives and conceptual implications", Psychological bulletin, 114(2), 296-322.
6 Sherrington, S. C. S. (1905), On reciprocal innervation of antagonistic muscles – Seventh-Eighth note (Harrison: London).

7 Berntson, G.G., and J.T. Cacioppo (2008), "The neuroevolution of motivation", in: Handbook of Motivation Science, eds. J. Shah and W. Gardner (New York: Guilford Press), pp. 188-200.
8 Berntson, G. G. and J.T. Cacioppo (2008), "The funtional neuroarchitecture of evaluative processes", in: Handbook of Approach and Avoidance Motivation, ed. A.J. Elliot (New York: Psychology Press), pp. 307-321.
9 Soon, C. S., M. Brass, H.-J. Heinze and J.-D. Haynes (2008), "Unconsciouss determinants of free decisions in the human brain", Nature Neuroscience, 11, pp. 543–545.
10 Shiffrin, R. M., and W. Schneider (1984), "Automatic and controlled processing revisited", Psychological review, 91(2), 269-276.
11 Edgley, S. A., J.A. Eyre, R.N. Lemon, and S. Miller (1997), "Comparison of activation of corticospinal neurons and spinal motor neurons by magnetic and electrical transcranial stimulation in the lumbosacral cord of the anaesthetized monkey", Brain: a journal of neurology, 120(5), 839-853.
12 Aron, A.R. (2008), "Progress in executive-function research: From tasks to functions to regions to networks", Current Directions in Psychological Science, 17, 124–129.
13 Ajzen, I. (1991), "The theory of planned behavior", Organizational behavior and human decision processes, 50(2), 179-211.
14 Petty, R. E. and J. T. Cacioppo (1986), Communication and persuasion: central and peripheral routes to attitude change (Springer-Verlag: New York).
15 Nosek, B. A., A.G. Greenwald, and M.R. Banaji (2007), The implicit association test at age 7: A methodological and conceptual review (Psychology Press: New York), pp. 265-292.
16 Phelps, E. A., K.J. O'Connor, W.A. Cunningham, et al. (2000), "Performance on indirect measures of race evaluation predicts amygdala activation", Journal of cognitive neuroscience, 12(5), 729-38.
17 Kircher, T., and A.S. David (2003), The self in neuroscience and psychiatry. (C.U.P.: Cambridge UK).
18 Pinker, S. (2003), "Language as an adaptation to the cognitive niche", in: M. Christiansen & S. Kirby (eds.), Language evolution: States of the Art (New York: Oxford Univ. Press).
19 Gazzaniga, M. S. (1995), "Principles of human brain organization derived from split-brain studies", Neuron, 14, pp. 217–228.
20 Cacioppo, J. T., and G.G. Berntson (1992), "Social psychological contributions to the decade of the brain: Doctrine of multilevel analysis", American Psychologist, 47, 1019-1028.
21 Dunbar, R. (2003), "Evolution of the Social Brain", Science, 302(5648), pp. 1160-1161.
22 Dunbar, R. I. M. (2009), "The social brain hypothesis and its implications for social evolution", Annals of human biology, 36(5), 562-572.
23 Dunbar, R. I., and S. Shultz (2007), "Evolution in the social brain", Science, 317(5843), 1344-7.
24 Chicago Social Brain Network (2011), Invisible forces and powerful beliefs: Gravity, gods, and minds (Upper Saddle River, NJ: FT Press).

25 Cacioppo, J. T., and B. Patrick (2008), Loneliness: Human nature and the need for social connection (New York: W. W. Norton & Company).
26 Cacioppo, J. T., and J. Decety (2011), "Social neuroscience: Challenges and opportunities in the study of complex behavior", Annals of the New York Academy of Sciences, 1224, 162-173.
27 Gregory, R. L. (1987), Oxford Companion to the Mind (O.U.P.: Oxford, UK).
28 Haggard, P., and M. Tsakiris (2009), "The experience of agency: Feelings, judgments, and responsibility", Current Directions in Psychological Science, 18, 242-246.
29 Herrmann, E., et al. (2007), "Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis", Science, 317, 1360-1366.
30 Nosek, B. A., A.G. Greenwald, and M.R. Banaji (2007), The implicit association test at age 7: A methodological and conceptual review (Psychology Press: New York), pp. 265-292.
31 Porter, R. (1987), "Functional studies of motor cortex", Ciba Foundation Symposium, 132, 83-97.
32 Povinelli, D. J., H.K. Perilloux, J.E. Reaux, and D.T. Bierschwale (1998), "Young and juvenile chimpanzees (Pan troglodytes) reactions to intentional versus accidental and inadvertent actions", Behavioural Processes, 42, 205-218.
33 Rescorla, R. A. & A. R. Wagner (1972), "A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement", in: A.H. Black & W.F. Prokasy (eds.), Classical Conditioning II: Current Research and Theory (New York: Appleton-Century-Crofts).
34 Rinn, W. E. (1984), "The neuropsychology of facial expression: a review of the neurological and psychological mehanisms for producing facial expressions", Psychological bulletin, 95(1), 52-77.
35 Shultz, S., and R. I. M. Dunbar (2006), "Both social and ecological factors predict ungulate brain size", Proceedings of Biological sciences from The Royal Society, 273(1583), 207-215.
36 Silk, J.B., et al. (2003), "Social bonds of female baboons enhance infant survival", Science, 302, 1231-1234.

euresis journal

euresis
Association for the Promotion of the Scientific Endeavour

Volume 3
Summer 2012

journal homepage
www.euresisjournal.org

# Can we trust others in our pursuit of knowledge? Anthropological and theological reflections on belief and the neurosciences

Javier Prades

Faculty of Theology, San Damaso University, 28005 Madrid, Spain
jprades@sandamaso.es

## Abstract

*The brain-mind debate is a field where interdisciplinary research is needed more than ever. Indeed, it behooves us to ask the question: "Is it possible that non-conscious parts of matter produce consciousness?" The answer to this question is of great importance to both the personal and social spheres of living. When entering this debate, issues of trust and belief arise within scientific reflection and this, consequently, leads also to the ethical dimension of the scientific disciplines that seek to study the brain and its relation to mind and language. As a point of departure we state some fundamental criteria about the nature of the scientific theories, and about a proper understanding of human reason. In this way it will be possible that interdisciplinary debate contribute to a humanism open to the unity of knowledge.*

*We provide a brief summary of the history of the epistemological nature of trust and belief. In fact this debate has accompanied epistemology in both the Anglo-Saxon and Continental traditions. The roots of this debate are already present in the great classical tradition with its questions about the nature of true knowledge. More recently, the debate has been enriched precisely because of its connection with the different conceptions of the relationship between mind and brain. Our contribution – anthropological and theological in nature – wants to hold that a strict causal relationship does not necessarily follow from brain phenomena to mental phenomena (including belief), although we recognise a clear correlation between these two phenomena. We insist that recent researches show what Jürgen Habermas calls the insurmountable character (Nicht hintergehbarkeit) of the duality of both perspectives, which arises from a deeper unity. In the final analysis, this "dual unity" remains an open question.*

*Situated in this context we examine the phenomenon of trust and belief as common ground for interdisciplinary reflection. Particularly we suggest that an examination of the area of interpersonal relationships as a basic human experience has crucial implications for the ethical and epistemological questions that can ultimately serve to a better understanding of human reality in all its dimensions. We conclude by showing how the anthropology of the imago Dei, typical of Judeo-Christian revelation is able to understand this mysterious dual unity of man and contribute to the public debate about the interrelationship between mind, brain and language.*

## 1. Introduction

In a forum such as this one, dedicated to the natural as well as the social sciences, my own contribution could seem alien to the perspectives of the majority of the forum's presenters. After all, my presentation is a reflection of a philosophical and theological sort, founded upon some fundamental premises of Christian revelation.

My point of view is based on the Judeo-Christian idea of man as *imago Dei*, gathered from the biblical narrative (Gn 1, 26-27; 2, 18-25), and interpreted in the light of the New Testament (Col 1, 15-20; 1 Cor 15, 45ss; 2 Cor 4, 4). Starting from these scriptural premises, theology has developed the notion of a "dramatic anthropology", to use a phrase coined by Hans Urs von Balthasar. The Swiss theologian saw that the biblical conception of man as *imago Dei* could be translated into a "dramatic anthropology." With this formulation, von Balthasar points out a tension between unity and difference both in itself and as it manifests itself in three "polarities", or tensions, which are constitutive of the human being. These polarities do not annul the unity of the 'I', of the singular person who asks himself about himself, but they show that the person bears an inevitably "dramatic" character. When man observes himself in action he recognizes in himself a triple polarity or tension which is not reducible to either of the two poles: 'soul-body', 'man-woman', 'individual-community'. [1, 2, 3]. The human person therefore cannot be explained within a purely monistic or dualistic conception, rather the person experiences his own self as a mysterious "dual unity" within each of these polarities. If one were to explain each of these polarities, as well as the deep connection between the three of them, one would be describing, in my judgment, that "mystery of the unity of the self" which gives the title to our Symposium.

I will return to these questions at the end of my presentation. First it is necessary to clear the path of possible objections which could imply a preliminary exclusion of a presentation of this sort. Is it possible, instead, to find a common ground and a common task between the perspective that I am proposing and that perspective proper to a conference where the prevailing participation is that of specialists in the various scientific disciplines dealing with the problem of the mind, the brain, and language?

My conviction is that it is possible to find such a space of common interest, if we respect the two rules that Jürgen Habermas has been proposing for some time. The German philosopher is concerned that all the participants of pluralistic societies may offer their best contribution to political debate, so that the profound ethical and social challenges affecting Western democracies may be approached in an adequate way. This is why he asks, on the one hand, for what he calls the cognitive translation of the content of religious traditions into terms which are universally comprehensible and, on the other hand, for an overcoming of what he calls secular narrowness [29, 5]. Habermas thinks that, when both of these conditions are respected, the cultural, social, and moral debate would benefit from all of the cognitive

resources present in our society, without prior exclusion of anyone. The seriousness of the challenges that we have before us do not allow us to reject any reasonable contribution to society. In the end, we should be able to verify whether this Symposium has been able to fulfill Habermas' aspiration, around questions so fundamental for the adequate understanding of the person and social life as those which deal with the brain, the mind, and language.

## 2. The sciences and trust

In order to approach our task it would be useful to remember the words of David Baltimore, given in the Whitehead Institute a few years ago. This Nobel laureate maintained that trust is a foundational element in academic life and that from it opens up an interdisciplinary horizon of greater interest, precisely for the task of understanding more deeply the significance of trust in academic and scientific life as well as in social life[1]. His words are an example that shows how the scientific world is able to assume the importance of trust between people as a decisive factor for the progress of rational knowledge. We have then a point of encounter in the category of trust because trust turns out to be decisive for scientific knowledge as well as for the philosophical (and theological) understanding of knowledge. According to Baltimore, the neurosciences are called to a deeper understanding of this human experience from their own point of view, while also assuming an interdisciplinary perspective which would allow for the establishment of better relations between knowledge and trust. Wittgenstein himself has already made a claim about the primordial and original role which trust plays in the search for knowledge [6].

If this is so, then nothing impedes the extension of our reflection to also cover belief, which is closely related to trust. In this sense, one may value all the legitimate manifestations of belief, which come from a primarily intersubjective meaning, from the sense of trust which some men place in others within many different dimensions of social life[2], until it reaches religious meaning, as a phenomenon typical of human societies, and then finally becomes a trust which is specifically Christian (the theological virtue of faith as we understand it). It is in one of those areas where the twofold condition formulated by Habermas (cognitive translation of religious traditions and the overcoming of secular narrowness) could yield greater fruit.

But there is still one more point of encounter between the world of experimental science and the world of philosophy and theology. I am referring to the numerous voices who from the world of the neurosciences are asking strictly ethical questions, which are also very much related to trust and with belief in other people. Indeed, a few years ago, the term "neuroethics" was coined in order to highlight the fact that scientists themselves feel the need

[1] See the transcript of this conference at: www.wi.mit.edu/news/archives/2002/db-0219.html
[2] In the field of sociology today we can find interesting studies on basic trust and trust in complex systems. Cf. [9].

to reflect about the moral implications of their techno-scientific discoveries[3]. In this way, together with the question about trust and belief, we find ourselves also with the question about the ethical dimension of science and technology. If the first question is epistemological in nature, concerning itself with the relationship between trust and knowledge, the second orients itself directly toward the anthropological and moral perspective. On both questions it appears legitimate to assert that the scientific discussion does not preventively close itself off before an interdisciplinary reflection about trust/belief, which carries with it the dimension of intersubjectivity. On the contrary, one rather comes to the conclusion that these are open questions in the world of science, about which a deeper understanding is desired.

It is possible then to discover a space for an encounter between different epistemological perspectives, when the time comes for us to approach the main question of this Symposium, which we could formulate thus: Is it possible for consciousness to arise from non-conscious matter? This is quite a problem, that of explaining consciousness, and for many researchers it "is the most complicated problem on the table for science[4]" [8].

Because I am not a specialist in any of the experimental sciences, I will quickly clarify that I will avoid entering into the technical questions within those fields in which many of those present here are world-renowned experts. Rather, working off the information available to a person of average intelligence in our Western world, I will try — as indicated from the start — offer my theological and philosophical perspective, illuminated by Christian revelation.

## 3. A few principles concerning the use of reason and the sciences

Before entering into theological questions, I think it is necessary to remember a few principles which govern the conversation between the scientific point of view and the philosophical-theological point of view. I am referring in particular to three questions which are closely related to each other: *Interdisciplinarity*, the necessary *difference and connection* between scientific and philosophical knowledge, and the inevitable demand for a *unitary point of view* with regard to the human person.

[3] About the birth and status of *neuroethics*, see: www.unav.es/cryf/neuroetica.html

[4] This difficulty is recognized by many scientists and philosophers. These words from Ph. Clayton serve as a good example: "Nobody has the slightest idea how anything material could be conscious. Nobody even knows what it would be like to have the slightest idea about how anything material could be conscious. So much for the philosophy of consciousness. Given the difficulty of the transition from brain states to consciousness, one might worry with Colin McGinn that we face here an irresolvable mystery" [10]. Also J. Searle: "How, for example, can it be the case that the world contains nothing but unconscious physical particles, and yet it also contains consciousness? How can a mechanical universe contain intentionalistic human beings – that is, that human beings can represent the world to themselves? How, in short, can an essentially meaningless world contain meanings?" [12]. Concerning Searle's position see the critical approach of R. Tallis [13]. D. Chalmers maintains that when investigating consciousness, one must point out the "easy problems": questions which deal with the functioning of memory, of learning and others referred to the question of how the cognitive function is realized, and the "hard problem" of knowledge: how do physical processes give rise to consciousness? [14].

### *3.1 Interdisciplinarity*

We can affirm without controversy, as a statement of fact, that *interdisciplinarity* lies at the very origin of the neurosciences, from its first steps up to its establishment as a discipline towards the end of the twentieth century [11]. Today, neuroscience is "the experimental science which has searched the most for a relationship with other disciplines, finding itself with questions which cannot be resolved exclusively with its own experimental methodology." [11]. The attempt to offer clear answers to the question about the global operation of the brain and its role in the complete life of the human person drives many scientists to look for such a collaboration, just as we see in the very hopeful activity of many university departments.

In this interdisciplinary effort, we should be conscious of the fact that the progress of understanding in these matters is a slow and limited, but reliable assemblage of new models which allow us to advance in response to very complicated questions. It is not an easy task or one that can be finished comprehensively. In his study of the reality of the "I", Antonio Millán Puelles maintains that "to be intelligible is not the same thing as to be easily intelligible, nor is it the same thing as being completely intelligible[5]" [15].

On the other hand, this interdisciplinary opening is very interesting because "the context of a possible positivism could condition the actual development" [11] of the neurosciences. Instead, the fact of interdisciplinarity could effectively help to correct the danger of this positivism, of which some authors warn us when we enter into the realm of neuroethics: "The good news about neuroethics was that experimental science had become conscious of its limits and appealed to a dialogue with other disciplines, even the philosophical ones. The bad news, on the other hand, is the fact that researchers in neuroscience have later on opted not for dialogue, but rather for a unilateral scientific explication of moral questions. Here is the great ambiguity that comes when one tries to evaluate the rise of neuroethics[6]." It seems necessary, remembering Habermas's thesis, to incorporate interdisciplinary criteria into our discussion if we desire that the great anthropological and ethical questions which have been raised by neuroscience be not diminished nor reduced.

### *3.2 Scientific knowledge within the matrix of human knowledge*

As you will expect, I do not pretend to explain to scientists what their work consists of. My intention is rather to point to a problem which is complex but decisive for our objectives,

[5] S.L. Jaki also appreciates this notion of the slow conquest of man's knowledge over man. "Unlike an angel who needs no conquest, and unlike an ape uninterested in them, man thrives on conquests which are the fruit of a mysterious union in him of matter and mind"[17].

[6] S. Sánchez-Migallón, *"La ambigüedad de la neuroética"* in: www.bioeticacs.org/?dst=neuroetica.

namely the place that scientific knowledge occupies within the matrix of knowledge which, as human beings, we are able to acquire about reality. From such a large question, I will limit myself to recalling a few basic affirmations.

It is well known that the excesses of philosophical idealism had the effect of favoring the appearance of positivism which, as a reaction to such idealism, reclaimed a necessary autonomy for the experimental sciences. However, that positivism swung the pendulum toward the opposite extreme of philosophical excess, according to which the scope of knowledge was reduced to what could be verified through empirical experimentation. This is the viewpoint which we can call "scientism." For Habermas, "'scientism' means the faith in science as such, or to put it differently, the conviction that we can no longer understand science as one possible form of knowledge, but rather that we should identify knowledge with science." [16]. Scientism is therefore that epistemological position in which experimental reason is the only proper mode of the exercise of reason, ignoring that human reason can be exercised in other modes which are irreducible to the purely experimental mode. There is no want of authors who warn us that a positivistic epistemology carries out an illegitimate reduction of the integral nature of human knowledge. John C. Polkinghorne wrote that "[…] there are questions which arise from science and which insistently demand an answer, but which by their very character transcend that of which science itself is competent to speak. There is a widespread feeling among practicing scientists, particularly those of us who have worked in fundamental physics that there is more to the physical world than has met the scientific eye." [18]. Ludwig Wittgenstein makes a similar warning: "We feel that even if all possible scientific questions be answered, the problems of life have still not been touched at all." [19].

The explanation of the process which took place within positivism would require time and competence which I lack, but it suffices to recall the well-known Vienna Circle manifesto (1929) which epitomizes the positivistic point of view and its materialistic consequences[7]. For our purposes we must point out that positivism illegitimately censors questions which are proper to human reason. As George Steiner puts it, "The positivistic postulate, according to which an adult conscience would only ask the 'How?' and never the 'Why?' of the world and of existence, is an act of censorship of the most obscurantist sort. It muzzles the voice which lies beneath the voices that are within us." [20]. But there is also the fact that positivism cannot help but fall into the contradiction that is rationally impossible to identify experimental knowledge with knowledge *tout court*, excluding other forms of knowledge. The contradiction becomes clear once one declares, without empirical verification, that a proposition is meaningful only when it can be the object of empirical verification[8].

Looked at in a different way, what this contradiction brings to light is the reflexive capacity of human intelligence, which always exceeds merely experimental activity. When one

[7] About the birth of the Vienna Circle and its scientific and ideological principles, see [23].
[8] L. Kolakowski defines the scientistic mentality as "the irrationality of positivist rationalism" [24].

commits this positivistic reduction, which we call scientism, one makes the claim that only the scientific method is rational. The consequence, in that case, is that scientific knowledge could not be limited by any other knowledge outside the scientific theories themselves and would be purely self-referential [21, 22]. Therefore, the scope of reality itself would be reduced, because there could not be, or would not be, a reality outside of that which is known through experimental methods. The result is what Herbert Marcuse has called a "synthetically impoverished world[9]." If this reduction always goes against the reasonable exercise of reason itself in any area of reality, its prejudices would be even more grave when we are dealing with — as is the case in neuroscience — a reality specific to human consciousness. For this reason, in both the epistemological order and the ontological order, in the plane of knowledge and in the plane of the extension and depth of reality, experimental science *must* enter into an interdisciplinary dialogue with other forms of knowledge like philosophy and theology, which provide knowledge from other perspectives which are just as legitimate and irreducible as the scientific perspective[10].

In the twentieth century, many philosophers have claimed the irreducible character of different areas of knowledge, as well as the necessity of mutual collaboration between them. It is enough to remember the names of Husserl, Habermas, Gadamer, Merleau-Ponty, and Ricoeur as examples of thinkers who have made the effort to rethink the irreducible character of everyday knowledge and that knowledge which today we group under the term, "sciences of the spirit[11]." Both such levels of knowledge are in some way presupposed in the exercise of scientific knowledge[12]. In particular it is philosophical reasoning which asks

[9] *"El descuido de la dimensión filosófica específica ha llevado al positivismo contemporáneo a moverse en un mundo sintéticamente empobrecido […] y a crear más problemas ilusorios de los que ha destruido […] Una falsa conciencia mutilada es colocada como la verdadera conciencia que decide sobre el sentido y la expresión de aquello que es. El resto es denunciado —- y endosado —- como ficción o mitología"* [25].

[10] R. Spaemann reminds us, ironically, that several sciences have had the pretension of erecting a total explanation of reality and that, today, looking back, these claims have turned out to be absurd. He warns that what has happened to psychology and sociology in the last century could happen to the neurosciences today [30].

[11] Let us not forget that the philosophical work of E. Husserl begins with the overcoming of a purely "psychologistic" (today we could say "naturalistic") understanding of logic. He demonstrates the irreducibility of essences or ideas to the material processes which support them. Cf. [31]. Spaemann employs Husserl's work in order to defend his claim that meanings will never be located in neurons: *"Intentionale Gehalte haben keinen Entsprechung im Gehirn. Sonst müssten wir die Infinitesimalrechnung ebenso in einem Gehirn ablesen können, wie ein Streichquartett von Mozart oder Michelangelos Pietà Rondanini. Die Hirnsforschung wäre die Integrationswissenschaft für alle Natur- und Geisteswissenschaften. Und ein Fehler in einem mathematischen Beweis wäre genau so ein positiver Hirnzustand wie die Korrektur dieses Fehlers, ein Zustand, der letzten Endes aufgrund physikalischer Gesetze in einem anderen übergeht. Man muss diese Dinge nur beim Namen nennen, um ihre Absurdität zu erkennen"* [30]. See also [13].

[12] For M. Merleau-Ponty scientific knowledge about the human person must be related to his lived experience of the world, because scientific knowledge is relevant only insofar as it is situated in a wider scope: *"Je ne suis pas le rèsultat ou l'entrecroisement des multiples causalités qui déterminent mon corps ou mon 'psychisme', je ne puis pas me penser comme une partie du monde, comme le simple objet de la biologie, de la psychologie et de la sociologie, ni fermer sur moi l'univers de la science. Tout ce que je sais du monde, même par science je le sais à partir d'une vue mienne ou d'une expèrience du monde sans laquelle les symboles de la science ne voudraient rien dire. Tout l'univers de la science est construit sur le monde vécu et si nous voulons penser la science elle-même avec rigueur, en apprécier exactement le sens et la portée, il nous faut réveiller d'abord cette expérience du monde dont elle est l'expression seconde. La science n'a pas et n'aura jamais le même sens d'être que le monde perçu pour la simple*

about the place of science within the matrix of human knowledge, as well as the presupposed conditions for the possibility of an experimental science which offers a trustworthy understanding of reality. Let us consider, for example, the essential importance of notions like *demonstration*, *certainty*, *truth*, *coherence*, *nature*, *rationality*, etc., which are part of a philosophical theory of knowledge and without which experimental knowledge would be impossible. We can classify these assumptions under three categories: *anthropological* assumptions, which refer to science as a human activity; *epistemological* assumptions, which situate science within the context of rationality with its logical and gnoseological resources; and *ontological* assumptions which give order to natural reality within which science develops its great contributions [26, 27, 28, 61].

When, in the name of experimental science, the attempt is made to negate or ignore these assumptions, and with that, to deny that human reason is wider in scope than strictly scientific reasoning, or that reality is greater than the particular field on which an experiment is concentrated, then we find ourselves, properly speaking, not before a scientific theory, but before a (bad) philosophy of science, which Habermas calls "scientism[13]." And the reason for this is the same that we argued about positivism: it is an attempt that is not empirically verifiable to exclude all knowledge which is not empirically verifiable, such as meta-scientific knowledge.

### *3.3 The inevitable need for a unitary view of reality*

What becomes manifest in this attempt to exclude non-scientific knowledge from the realm of real knowledge? If we think about it, we see that the existence of scientistic discourse is due to the fact that the scientist cannot help but to elaborate a comprehensive theory of knowledge, which exceeds the scope of his scientific discoveries, precisely by virtue of a need proper to his rational condition. In this sense, scientism is, in spite of itself, a testament to the existence of a rational attitude and an inclination toward truth within every thinker. If the scientistic argument exists it is because of the admirable fecundity of scientific knowledge, which projects itself beyond experimental data and the theories which explain that data, so as to become open to the need that the human person has to comprehensively interpret his own existence. This need is also related to the need that reason has to open itself up to the totality of the real. In this sense, scientism and all positivistic claims contradict themselves when they become rational theories—not scientific theories—going against their own fundamental claim, which is the reduction of knowledge to purely experimental knowledge.

*raison qu'elle en est une détermination ou une explication"* [32]. Indeed, a scientist always, whether he knows it or not, thinks from within a worldview (*Weltanschauung*), as Popper, Kuhn, Lakatos and Toulmin have all argued, just to cite a few. Put differently, *"toda actividad científica incluye necesariamente una precomprensión no justificable científicamente"* [33].

[13] *"Der szientistische Glaube an eine Wissenschaft, die eines Tages das personale Selbstverständnis durch eine objektivierende Selbstbeschreibung nicht nur ergänzt, sondern ablöst, ist nicht Wissenschaft, sondern schlechte Philosophie"* [35].

But there are more than just, let us call them "negative" theories. The history of scientific reflection in the twentieth century offers many splendid examples of thinkers who have attempted to reconcile experimental knowledge with a unitary conception of man and reality. We can think, from the point of view of physics, of the well-known and very interesting discussions about the relationship between science, ethics, and religious faith which took place in the late 1920s between Dirac and Pauli, or Planck and Einstein, as they are recollected by Heisenberg[14].

We can also remember the great effort made in the same interwar period by thinkers like Gehlen, Plessner, or Scheler in anthropobiology and philosophical anthropology. Through the study of biology and animal ethology, they began to understand the specificity of human corporality, which is linked to the human person's opening towards reality[15]. Max Scheler's distinction between the *Umwelt* (environment) proper to animals and the *Welt* (world) proper to the human person makes it possible for the latter not to be enclosed within the strict chain of needs and satisfactions [34]. The human person can take a distance from himself and can and should understand himself, in a manner which reaches a conception of himself, of what he is and what he should become. This path of inquiry has been taken by more than a few phenomenologists who have known how to describe the properly human condition of corporeality. Arnold Gehlen claims that the biological indeterminacy of the human person—his corporeal expressiveness—corresponds to the human intelligence's constitutive opening towards the totality of the real [39]. As we have already seen, by the 1920s, experimental investigations were already able to claim that human corporeality is specifically human, correlated with an intelligence and consciousness open to the totality of the real and an adequate knowledge of itself.

Our thesis is that the need to know one's self, which is alluded in the title of this Symposium, is the point of encounter between ordinary knowledge, scientific knowledge, philosophical knowledge, as well as artistic and cultural and even the ethical and religious dimensions upon which we build our lives. Helmuth Plessner defines this state of affairs unique to the human person among the other animals as an "eccentric" position by virtue of which the human person does not only have a *surrounding world*, but a *world*, strictly speaking [40]. The human person sees water as a response to thirst, but also knows how to see it as $H_2O$, and knows how to see its beauty when it runs through rapidly or falls in a waterfall, knows how to see it as a symbol of moral and religious purification, etc. Put in another way, the human person is able to recognize the truth of a mathematical formula or the coherence of a scientific explanation, is capable of leaving himself and reaching out towards reality itself and, thanks to the uniqueness of its own corporeal constitution, is capable of possessing himself when he goes out in search of another. This is a way of describing that immense

[14]See the presentation on this topic by W. Heisenberg [36]. Cited and commented by J. Ratzinger [37].
[15]For a study of this problem, which we cannot deal with in this paper, see [38].

depth which we are referring to when we speak of self-consciousness, a subject about which contemporary neuroscience has greatly contributed to our understanding. When one renounces the convergence of these perspectives, that is, when one renounces the world, for a scientistic reduction, the human spirit finds itself confined to a narrowness which does not correspond to its demands, to the point in which it finds itself in anguish. The human person needs —- above all reductionisms —- a unitary vision of the world and of himself which frees himself from existential anguish [41, 28, 15]. Given that the person goes beyond the environment (*Umwelt*) and opens himself up to the world (*Welt*), it is necessary to understand consciousness itself in a way which does justice to that multidimensional scope to which we have alluded. Habermas has referred to the variety of modes of knowledge ("biological" and "social" in type) and has with respect to them recalled the "human need of a unitary vision of the world." [5].

When the sciences, philosophy, and theology share a unitary perspective it will be possible that all of us contribute to the decisive task of developing a humanism which accepts the possibility of a unity of knowledge, decisive for university life and for the social life of our western democracies[16]. In particular, philosophy and theology will fulfill the important social task of helping scientific knowledge not to fall into the reductionsisms that I described.

It is within this context that all of the reflections which we would like to offer about trust and belief find their significance. It does not seem difficult to concede that trust is essentially related to a being whose position is naturally eccentric, whose intelligence is open to the totality of the real and whose consciousness is capable of taking a distance with respect to itself. It is precisely in a being with these characteristics that the trust in another is not only possible but also strictly necessary. For this reason, we think it is legitimate to proceed with our investigation about the relationship between trust and certainty/faith.

## 4. The debate about knowledge and trust in the history of philosophy

The problem of the influence (positive or negative) of belief over knowledge is not limited to the sciences but also forms a part of the history of western thought. We can recall briefly a few of the most significant episodes in that history.

The need to establish the conditions for true knowledge, its difference from apparent knowledge, and the distinction between *episteme* and *doxa* finds its roots in Greek philosophy.

[16]There is no shortage of voices that rise up against any effort to recuperate a unitary vision of knowledge. This rejection usually occurs when such an effort is identified with one of the forms of European idealism of the 19th century (usually, Hegel), whose political and social consequences were tragic in the 20th century. Indeed, any effort to construct a rational system in which the human person absolutizes himself cannot but have terrible, dehumanizing consequences. From there it will be decisive to think about the unity of knowledge while respecting the creaturely condition of the person and, moreover, his double characteristic of finitude and openness to the infinite.

From there it passes through the great authors of Christian antiquity, like Augustine, and enters into the Middle Ages (Thomas Aquinas). The classical world offers us a proper distinction between knowledge and belief which we should keep.

At the beginning of modernity, this distinction entered into a new phase characterized by a strong contraposition between knowledge and belief. The dominant currents of modern thought (Descartes-Kant in the continent and Locke-Hume in the Anglo-Saxon world, to cite the key authors) chose to relegate that type of knowledge dependent on trust in another (belief) to an inferior rank, never comparable to the true knowledge which comes from an immediate perception of itself (Descartes' *cogito*) or the direct experience of directly observable properties (empiricism). For both ways of thinking, the intervention of trust reduced — or eliminated — the epistemological value of knowledge. With a formula which came to be very widespread, "That which is known is not believed, and that which is believed is not known." We find the most radical version of this profound separation in positivism, which we have already spoken about, which delegitimizes any knowledge which is not strictly verifiable by experimental methods. A fortiori, trust and belief would be left out of the nucleus of knowledge and form part of the emotive or sentimental dimension of the human person.

However, in the twentieth century, the doors of the debate about forms of knowledge were re-opened, precisely because science went deeper into the human condition. The human person's social constitution is rediscovered, as it is reflected in language and in the social dimension of knowledge. When one goes through the literature concerning "witness" — which is a form of knowledge supported by trust in another — one often finds critiques of the dominant forms of knowledge in western thought. The accusations are epistemological in nature, and, as such, they critique subjectivism or naturalism for the risk that they may identify a person's knowledge with an object's knowledge. What justifies such denunciations is that one may consider those to be unacceptable defects for reaching a complete vision of the human person, respectable of his dignity. If subjectivism and naturalism deserve reproach, it is because it is judged that they are insufficient for reaching the heights of the human condition.

There are more than a few authors who justify knowledge-by-witness while critiquing modern epistemology. Tony Coady warns that forms of knowledge dominant in the West are individualistic[17]. On his part, Claude Bruaire critiques the autosufficiency of a thought which has desired to conquer the world while betting everything on scientific and technical reason[18]. Giuseppe Angelini denounces subjectivism, both ancient and modern, as well as naturalism, and criticizes the "reciprocal estrangement" between men in the Hegelian thesis

[17] "In the post-Renaissance Western world the dominance of an individualist ideology has had a lot to do with the feeling that testimony has little or no epistemic importance" [43]. See also recently J. Lackey [44].

[18] *"L'humanisme et sa volonté d'autosuffisance a voulu conquérir notre monde et se saisir de notre destin en misant tout sur la puissance d'une raison scientifique, capable de la plus grande efficacité technique"* [45].

that civil society is a system of needs[19].

Beyond denouncing the "individualistic" reductions, literature also picks up on "relational" or "communitarian" forms of knowledge. Throughout the twentieth century, there have been philosophical trends which have sought to recuperate the relational or dialogical dimension in the definition of the human person [43, 7]. The social character of knowledge is developped both by the rise of philosophy of language in Anglo-Saxon thought, and by continental thought[20].

As we can see, there are many trends which today insist on conceding importance to the social and communitarian dimension of knowledge. A thinker who adopts this perspective, such as Habermas, rejects biological naturalism and attempts to overcome its reductionism precisely by appealing to the social condition of the human person[21]. His thought shows us that the problem of trust and the problem of neuroscience are not as far apart as it would seem at first glance. On the contrary, from within the debate about the relationship between the mind and brain — with all the different possible explanations — the need arises to relate this relationship with the relationship between individual consciousness and social consciousness. Both questions deal with two of the constitutive polarities which we spoke about above: "soul-body," and "individual-community." To approach these both separately and as related to one another are the decisive tasks for entering into questions of ethics and religion. Let us first see the problem concerning the relationship between mind and brain, and after that we return to the problem of trust.

## 5. Philosophy of mind and the relationship between mental and brain processes

Concerning the relationship between mind and brain, we will limit ourselves in this paper to affirm that the correlation between judgments and theoretical and practical decisions (i.e., mental processes) on the one hand, and neurobiological process on the other, do not necessarily imply a complete causal relationship of the latter on the former. *Post hoc* is not the

[19] *La difficoltà dipende dalla pressione esercitata dai luoghi comuni della cultura corrente. Mi riferisco in particolare, a un preciso luogo comune, che appare assai tenacemente iscritto in tutta la cultura dell'Occidente: mi riferisco ad una rappresentazione incautamente soggettivistica del soggetto"* [47, 48].

[20] *"La socialización de la cognición [es] característica del espíritu humano"* [5], and also [42, 49, 50, 34].

[21] The German philosopher rejects the naturalistic reduction of knowledge and freedom, posed from the point of view of neuroscience, and reclaims the irreducible character of knowledge regarding its biological infrastructure, as well as its "spiritual" condition. His argument consists precisely in pitting against biologism an evolutional understanding of knowledge which is intersubjective and social in nature. He argues that a "mentalistic" vocabulary cannot be completely translated into a "biologistic" one without losing some certain aspects of the human person. One must add, however, that his final proposal does not seem to be sufficient, because even if he does not argue that the "I" is reducible to a pure social construction, practically speaking he identifies it with a linguistic structure which makes social action possible. That "I" is nothing more than an integral part of a system of pronouns without any privileged position. Cf. Habermas [5].

same *as propter hoc*[22] [2]. The recent literature, of interdisciplinary scope, highlights this fact, as we shall see in this brief summary of the main theories.

### *5.1 Main theories about the relationship between mind and body*

From the middle of the twentieth century, philosophy of mind has concerned itself with the relationship between the mind (soul) and brain (body)[23] [51]. This new philosophical discipline assumes the task of reflecting on the formal nature of mental phenomena, overcoming unilateral positions and taking into account the contributions made by the cognitive sciences[24]. As Pascual Martinez-Freire says, it is necessary that "philosophy of knowledge develop taking into account the theses of the cognitive sciences" [52], among them, those of psychology, neurophysiology, linguistics, artificial intelligence, etc.

What we are proposing then is a complementary view to those we have been examining. We said before that scientific reflection cannot avoid orienting itself towards a unitary view of knowledge, and thus to open itself to philosophical and theological horizons. We add now that philosophical and theological knowledge cannot progress if it does not take into account the contributions of cognitive science. It is another reason for supporting the interdisciplinary bond between the themes of our Symposium.

While it might sound a bit simplistic, it could be enough to classify all theories about the relationship between the mind and the brain into three main groups[25]. The first we can call *materialist*, and within this group we can include theories like that of physicalist identity or monism[26], epiphenomenalism[27], and emergent monism[28]. In this type of theory, what

[22] Cf. Scola, *'Anima e neuroscienze'*. Other contributions can be found in the following interdisciplinary studies [53, 42, 30, 54].

[23] Philosophy of mind is a recuperation and updating of the philosophical-theological problem of the relationship between the soul and the body. We say "recuperation" because it has been a question dealt with in the West by the Greeks and later by the Christian tradition, but it lost relevance to philosophy—though not to theology—after the 18th century. We say "updating" because its reflections on the nature of the mental take into account findings made by cognitive science.

[24] Cf. C. Beorlegui, *'Filosofía de la mente. Visión panorámica y situación actual'* in: www.uca.edu.sv/ facultad/chn/c1170/Filosofia%20de%20la%20mente.pdf, 1-3.

[25] Information about these theories can be found in [33].

[26] A representative of this current of thought is H. Feigl. His central thesis can be summarized in this way: a) The mind and mental states are objective realities; b) The mind is the brain; c) The brain is, upon final analysis, a physical reality. Cf. [57, 58].

[27] This form of materialism *"sostiene que los fenómenos mentales existen, pero no son causalmente efectivos. Las propiedades mentales acompañan a los sucesos neuronales, pero no influyen sobre ellos. […] La conexión causal existe sólo en una dirección, la que lleva de los sucesos físicos a los mentales (o a otros sucesos físicos), nunca desde éstos a aquéllos"* [16].

[28] M. Bunge defends a form of emergent monism (as a fruit of evolution) or a systematic theory of the brain, consistent with a monistic understanding of substances (a physical-chemical brain) and a dualism of properties (the brain possess physical realities, but also mental ones). But this emergentism results in being a materialism because for Bunge there is nothing spiritual in the material and in the mental; all mental activities are always and only properties of the physical chemical brain. Cf. his work [59]. Ruiz de la Peña [58] contains a critique of Bunge's view.

we call mental processes and mental states are considered to be more or less sophisticated processes and states within the complex physical system of the human brain. If we were to summarize this type of theory in one maxim, it would be the phrase used by Francis Crick: "You are nothing more than a pack of neurons"[29] [55].

The second group, on the contrary, stands for a dualism whose roots one could find in Descartes. Within this group we can name, for example, the interactionist dualism of Karl Popper and John Eccles[30]: for these authors, mental states constitute a specific type of natural phenomenon which is essentially non-physical. Among the neuroscientists who defend mind-brain dualism we can also name Wilder Penfield, who claims the existence of a center of mental decision-making, distinct from the cerebral-neuronal framework, in the way that a "telephone operator controls a switchboard" [56]. We should also say that the dualist thesis is today rejected by the majority of philosophers and scientists. Effectively, dualism attempts to highlight an ontological, ethical, and spiritual peculiarity proper to the human being, but it juxtaposes two planes of reality without avoiding the grave risk of turning each into independent entities.

On top of these clearly monistic or dualistic theories, there exists something we can call the third way, in which we find philosophical and scientific theories as diverse as functionalism[31], emergentism[32], and dynamic structurism[33], to name a few. Despite important differences among them, they generally coincide with referring to the human person as a unitary reality essentially constituted by two inseparable moments: the mind (for some, the soul), and the body. They differ among themselves in the way of explaining this "dual unity" of the person.

[29]He claims that in a not-too-distant future we will be able to explain the pseudo-phenomenon of consciousness by appealing only to neuronal correlates. In the Spanish scene, a neuroscientific defender of the theory of identity is F. Mora, who affirms that *"la actividad cerebral son los procesos mentales"* [60].

[30]They explain their theory in [18]. In it, Eccles argues that *"la mente autoconsciente es una entidad independiente que se halla activamente entregada a interpretar la multitud de centros activos de los módulos de las áreas de relación del hemisferio cerebral dominante"* [18]. An explanation and critique of this theory is offered by [57, 58]. These writers consider Popper to be more of an emergent than a dualist. They are partly correct, because Popper was more moderate than Eccles in his stance, given that he defends the idea of consciousness-brain interaction, wherein mental phenomena exercise causal influence over the brain. Eccles explicitly defends the view that the soul is created directly by God in [62].

[31]Functionalism is born with the attempt to overcome Cartesian dualism as well as its opponents, behaviorism and the monistic theory of identity. It was first formulated by [64, 65]. In this way of thinking, one thing is the real physical support of thought, and another thing is the mental states themselves, which could be reduced to the brain. According to functionalism, what defines a mental state is the complex of causal relations that are maintained with 1) the environmental effects on the body and 2) other types of mental states, and 3) the body's conduct. For a synthetic exposition of the limits of functionalism, cf. [4].

[32]Emergentism takes many forms. Its central thesis is that the mind emerges from the brain; what the different forms disagree on is the different ways of explaining how that emergence happens, as well as the relation between mind and brain. One notable defender of neuropsychological emergentism is P. Sperry. He claims that mental states are emergent properties (of a higher rank) which come from the brain. Without falling into a dualism, he affirms that mental states do not happen independently of physical events, and defends that those mental states and those physical events are two distinct type of realities: *"las cualidades subjetivas son […] de índole muy distinta a la de las neuronas, moléculas y otros componentes materiales que les sirven de base"* [79]. More recently, emergentism has been defined as a claim (called "non reductive physicalism" or "monism with a dual aspect") by the neuropsychologists M. Jeeves and W. S. Brown, in their work [66].

[33]This is the position proposed by [67].

They admit that there is a certain rupture of continuity between the mind and the brain: although they do not do away with matter, they assert that the mental is something more than material. As we have already said, to explain the ultimate reason for this mysterious but real unity — that is, the relationship (partially causal or not causal) between mind and brain (or between cerebral mechanisms and mental processes) — is a very complex issue[34].

Whatever the case may be, and always having in mind that the diversity of theories within this third group, it is possible to prudently affirm that this third way between monism and dualism is the most compatible with the Aristotelian-Thomistic idea of the *anima forma corporis* or *anima forma materiae primae*, though with a complexity derived from the incorporation of data from the cognitive sciences. About this we will deal with below.

### *5.2 Philosophy of mind before the reality of free self-consciousness and its spiritual nature*

The root and, at the same time, most mysterious problem of the human condition is the fact of free self-consciousness. We can only approach this problem from an interdisciplinary perspective, which would be capable of taking into account neuroscientific data and at the same time be open philosophically and theologically to what is essential about subjectivity. Gunter Rager affirms the need for this interdisciplinary approach when he considers it necessary to equally value the understanding of consciousness that comes from the sciences as well as that which comes from "the world of life" (*Lebenswelt*) [54].

From its perspective, neuroscience can establish that, for example, when a man falls in love, a zone in the brain is activated which is different from the zone that is activated when he is having ice cream. It could also establish that, when one thinks introspectively, some neurons are activated, and some are not — just as, in a similar way, the toes on the feet do not move while he is in the process of introspection. But neuroscience cannot describe what self-consciousness formally consists of, nor can it localize consciousness, because consciousness does not occupy a place, even though it may emerge — necessarily, but not only —- from the brain[35]. Francisco Varela has conceived of a useful formula when he argues that the synchronicity of the brain and its dynamic operation are the "conditions of possibility" for the appearance of consciousness. But they are not sufficient conditions[36] [68].

[34] W. Penfield maintained that, as a neurophysiologist, it would never be possible to explain mental processes through the action of neurons in the brain. This is why he spoke about the brain as a mysterious thing.

[35] M. Kurthen summarizes in this way what neuroscience can (and wants) to do with regard to self-consciousness (*Selbst*) in relation with philosophy of mind: to propose to other scientific disciplines, or to common psychology, a concept of self-consciousness; to determine the relation between existential aspects of self-consciousness and their cerebral correlates; to clarify and value their premises in philosophy and theory of science; to judge whether self-consciousness can be explained neurologically [69].

[36] *"Las sincronías cerebrales son esenciales, son una condición de posibilidad sine qua non [para el aparecer de la conciencia ...]. Sólo una vez que hemos cumplido esta condición de posibilidad que permite que puedan establecerse los ciclos de acoplamiento del cerebro con el cuerpo, del organismo con el mundo y del organismo con sus partes, puede emerger la conciencia. [...] Esto es, por cierto, un ejemplo notable de lo que se conoce como 'fenómenos emergentes' de la teoría de sistemas dinámicos: tenemos una serie de elementos locales (neuronas, cerebro, cuerpo, mundo físico)*

From this affirmation we can explore philosophically what neuroscience calls the "first-person perspective." This is only possible if we overcome the scientistic prejudice which considers that which is subjective to be relative and non-scientific. A neuroscientist like Varela defends the subjective realm, the "first person[37]", in this way: "Traditionally, first-person data ('I feel that…') has been labeled as non-trustworthy; it is subjective, and subjective is always synonymous with arbitrary and capricious. This is an absurd taboo, because lived experience, the subjective, is part of nature, and if it is, then it can be expressed, studied, and validated" [68]. It is true that the experimental sciences can know and explain the "how" of something, but not its ultimate reason, its ultimate "why." Instead, this is the task of philosophy, which is equipped to adequately study "first person" data, as well as theology.

When we enter into the strictly philosophical realm we find ourselves with the reality of "consciousness," or the reality of the "I." Within this reality we have to distinguish at least two levels: the consciousness-of, or the state of being conscious of something (which we call intentionality), as well as the consciousness-of-oneself (which can be given as a consciousness concomitant with an intentional act or as pure reflection). Above all else, with this last form of consciousness we maintain that its nature is non-material, or spiritual[38].

Concerning what exactly the spiritual character of consciousness consists of, the opinions vary. Many philosophers accept the definition of spirituality as the constitutive opening of the "I" to all other things that are also an "I", and in general to reality as such. It is the dimension of the human person which includes intelligence, freedom, affectivity, morality, etc. One interesting account of all this, from philosophy of mind, is the explanation proposed by Martinez-Freire [52]. For this author, there are three types of mental processes: mental-cerebral processes (physical) which are proper to all animals (including human beings) with greater or lesser degree of complexity; mental-physical processes (non-cerebral) which occur in some machines; and finally, non-physical mental processes, which occur only in human beings, and which are defined as spiritual. With respect to the first two types, we can speak of a certain "emergence" because, in the first case, mental processes emerge from brain neurons, and in the second, they emerge from the physical process of computation. In contrast, with respect to the third process "one can defend the strongest dualist stance, because it [the process] appears independently of neuronal processes and at the same time utilizes neuronal processes." And he concludes: "the notion of spirit refers to spiritual pro-

*que cuando se acoplan, cuando entran en relación, dan origen a un fenómeno que no es la mera suma de partes, sino una globalidad, distinta y unitaria, y que modifica el funcionar de los elementos sin ser reducible a ellos. Por mucho que busquemos este fenómeno global, no lo podremos encontrar en ninguna de las partes en particular, ya que justamente emerge de la interacción de todas ellas"*,pp. 249.

[37]Varela defends a "neuro-phenomenology" that is, an epistemological path which unites the third person and the first person perspectives. That is, a science which combines both neurology and the phenomenology of consciousness: the lived experiences of a subject, and his vital and witnessed experiences [68]. Habermas also considers it necessary to complement the third person perspective with that of the first person [5].

[38]A detailed account of the dynamisms of consciousness can be found in [70].

cesses, which ought to be extended as processes which imply a non-physical causation, which employ neuronal resources but exceed those same resources" [52]. Among these spiritual processes, we can count self-consciousness, free will or freedom, the search for the common good, or the search for a personal life-project, to give a few examples[39]. Because of this, this philosopher concludes that freedom constitutes a quality exclusively within human beings, more particular to them even than intelligence[40].

## 6. Tendencies in philosophy of mind which reinforce the role of trust in knowledge

Given that our theme is the role of trust in others in the pursuit of knowledge, it is not enough merely to have claimed an interpretation of the mind-brain relationship which allows us to claim a "spiritual" character about the mind. No doubt that this was a preliminary and necessary step, because if the mind could be reduced completely to the brain, we would lack the motivation to keep questing about the significance of trust. Rather, trust would be another effect caused by cerebral activity[41]. If, as we have seen, there is room for an interdisciplinary take on the matter that is not reductive, neither monistic nor dualistic, and we recognize the reality of "non-physical mental processes," then we may explore more in-depth the importance of trust in knowledge, as Baltimore has asked that we do, and Wittgenstein with him.

We can also affirm that in contemporary philosophy of mind, there are many currents of thought which are of interest to our objective. We will refer to them briefly, keeping in mind what I have already said in section 4.

In the first place, we must recall how philosophy has recuperated the social or communitarian dimension of human experience. We have already mentioned this. Both Anglo-American and continental philosophy have been able to overcome an individualistic conception of the human person, and of his way of knowing. In the continent, phenomenology made the first steps from the beginnings of the twentieth century. For example. Millán Puelles is able to take some of Husserl's work and convincingly show that the concept of subjectivity always includes a transcendent impulse towards another being[42] [70]. For this

[39]Martínez Freire distinguishes free volitions and simple volitions. These latter ones are determined by stimuli or conditions under which they are normally situated: for example, the desire to eat at three o'clock. Free volitions, on the other hand, are an act of freedom, for example when a man chooses to go on a hunger strike or to accept martyrdom.

[40]This is why he affirms the existence of animal intelligence and mechanical intelligence, but not animal freedom nor mechanical freedom. Cf. [52].

[41]About the aporia which arises with the negation of the sense of reason by reducing it to a pure biological fact, see my paper [71].

[42]He considers it anti-philosophical to ignore or deny this evidence. He finds support for this in E. Husserl, who reflects on this problem in the fifth C*artesian Meditation*, and whom he considers the philosopher who has been able to give the best account of the *alter ego* starting precisely form the experience of the *ego*.

reason, any treatment of the "I" cannot be done without including what the author calls "the plural of the I." [15]. Evidently, we are not arguing against the unity of the "I", because this is always an individual "I", alone and unique, that adds itself to others. Rather, we are speaking of "the evidence that the ego finds itself within its own life as an alter ego precisely of the same quality as the other I." [15]. If this is so, philosophy of mind will be in the condition to enrich the understanding of selfhood which we have considered as the point of encounter of science, philosophy, and theology. Along with the "first person" or subjective perspective (of phenomenology), and the "third person" or objective perspective (of neuroscience), philosophy of mind begins to consider as necessary the "second person" or intersubjective perspective of social interaction[43]. This anthropological and epistemological perspective offers a foundation from which to consider knowledge in which trust in another — as in, for example, knowledge-through-witness – has full legitimacy both in the scientific realm as well as in the judicial, moral, and religious realms.

This relationship, typical of the human person, and which opens it up to other subjects, is already affirmed by the philosophical currents which emphasize the constitutive corporeality of the human person. We have already seen how, in the twentieth century, anthropobiological studies like those of Plessner and Gehlen have brought to light the singular corporeal condition of the human person, which makes him out to be an "eccentric," turned away from himself, open to others and in need of a mental or spiritual dimension, radically distinct from the corporeality of other animals. The findings of those studies are in harmony with those of more recent philosophical research, as in the work of Martinez-Freire, who defends the need to take into account the corporeality of the subject. He considers the "incorporation" of the subject to be as evident a fact as that of the existence of mental processes[44], without seeing the need to equate or reduce the spirit and mind, and the mind and brain [52]. The neurologist Antonio Damasio categorically affirms that without a body, there can be no mind, taking a firm distance from any Cartesian way of seeing things[45] [75]. We find here an objective nexus between the polarity of "soul-body" and "individual-community": the opening of the spirit to the real which comes from a singular corporeality of the human person which is united to the intersubjective or communitarian dimension, which makes the person not only an irreducible individual, but also, at the same time, a social being.

The corporeal and intersubjective dimensions of the person, which decisively influence all modes of knowing, are reflected in another area of current interest, that of emotion and affection. With respect to this topic, it could be enough to point out that cognitive science does not claim that the emotions are irrational in nature, but rather that they play a sig-

[43] The dialogical nature of the human person has been recuperated, starting from the precedent of L. Feuerbach, by modern philosophers like M. Buber [72], P. Ricoeur [73], or X. Zubiri [74].

[44] *"No sólo mis procesos mentales (sensaciones, percepciones, creencias, inferencias, recuerdos, sentimientos y voliciones) me son evidentes, sino que también me resulta evidente que soy un sujeto incorporizado"* [52].

[45] Concerning Damasio's position, see [80].

nificant role in the workings of intelligence, with a clear cognitive function[46]. Today, what is called "affective neuroscience[47]" enjoys wide interest. In this field, which began in the 1980s, two well-known neuroscientists, Michael Gazzaniga [76, 77] and Damasio, defend the thesis that feeling is a moment within, or an integral component of, reason[48]. As we have pointed out continuously in this paper, this appreciation of affectivity and its role in knowledge is not an exclusive finding of neuroscience, but rather "the rigorous and serious dialogue between neuroscience and philosophy which is the only valid path with which we can avoid falling into triviality or mutual disqualification." [78].

## 7. The dual unity of soul-body and individual-community according to the anthropology of the *imago dei*

As we enter into the final part of our presentation, we should recall the question which gives our paper its title: C*an we trust others in our pursuit of knowledge?* In response to this question we have tried to offer epistemological criteria that would legitimize an interdisciplinary approach to the problem of knowledge and the human person's ability to know about himself and reality. Our claim here is that the demand for self-knowledge is the realm in which neuroscience, philosophy, and theology can find each other. We have also claimed the need to maintain unity among two dimensions: the corporeal or biological (i.e., the brain), and the mental and spiritual (i.e., the soul), pointing out distinct responses to the complex question of *how* these unities are able to subsist. We have also contended that man is not only an individual but is also, always and simultaneously, a social animal. From there we have maintained that human knowledge and understanding is a reality with an organic basis and, at the same time, a spiritual character, and that human knowledge is a reality that is both social and individual.

From the beginning of this exposition we have said that our point of view is born out of a Judeo-Christian anthropology, in which the human person is considered as an image of God. And we delayed until the very end an evaluation of that anthropology with respect to the problems we have dealt with. Now, finally, we can offer a few brief theological considerations which would complete our task, though obviously we have had, from the very beginning, a theological motivation.

[46]See the observations made by the neurologist J.E. Ledoux [81, 82].
[47]As this neurologist already wrote years ago [63].
[48]Let us consider the successful divulgation of D. Goleman [83, 84]. For a historical and appreciative account of the role of the emotions in cognitive science, see [78, 52, 87].

*7.1 The common good and humanism*

Before we begin, we should note that our reflections here do not aim primarily to make some sort of dialectic between "theistic" and "atheistic" positions. We have already seen that Habermas asked those participants in public debate to overcome a narrow secularism so that the religious perspective could be able to make the effort to translate itself into public speech. In that spirit, our goal in this presentation has been to claim, with reasoning accessible to everyone, that humanism favors the common good. So the true adversary to our position is *anti-humanism*.

With respect to the question about the relationship between mind and body, the three positions we outlined could really be reduced to two, because strict dualism is rejected by the majority of philosophers and scientists. Therefore we find ourselves with two alternative positions: materialistic physicalism (monism), and the "dual unity" which affirms the unity of the person in the duality of essential or intrinsic dimensions: the mental, or ensouled, and the corporeal, or cerebral.

In our judgment, materialistic monism, as a philosophical theory —- leaving aside the objective value of scientific findings —- runs the grave risk of developing a theoretical anti-humanism. Indeed, it reduces the person to a mere organic machine, or a simple animal, fruit of a random evolution which is only quantitatively (but not essentially) superior to inferior animals, or — to be more explicit — to a mere "pack of neurons." We use the term "anti-humanism" precisely to signify that this position rejects the uniqueness of the human being, claims that his dignity and his personal values are relative and not absolute and, in final instance, rejects the human person's ultimate, mysterious meaning. This position considers categories such as "person," "I", "freedom", or "transcendence" as concepts which are empty because they are not scientific[49]. Nietzsche's words had anticipated one hundred years ago the consequences of today's naturalistic determinism: we abolish all human responsibility because free will does not exist[50].

The other position still in play, that of a dual unity, posits a theoretical humanism, which from the interdisciplinary point of view opens itself up to scientific as well as philosophical and religious knowledge. From it we can derive certain ethical consequences: theoretical respect for the uniqueness of the person and for the human in general, for his essential particularity, for his absolute value, and for his ultimate mystery. This position humbly affirms that there is a limit to the scientific understanding of that mysterious relation between mind

[49] C.S. Lewis' book on the risk of the of abolition of man is still relevant here [88]. See also [92, 89].

[50] *"Nun entedeckte man schließlich, daß auch dieses Wesen nicht verantwortlich sein kann, insofern es ganz und gar notwendige Folge ist und aus den Elementen und Einflüssen vergangener und gegenwärtiger Dinge konkretisiert: also daß der Mensch für nichts verantwortlich zu machen ist, weder für sein Wesen, noch seine Motive, noch seine Handlungen, noch seine Wirkungen. Damit ist man zur Erkenntnis gelangt, daß die Geschichte der moralischen Empfindungen die Geschichte eines Irrtums, des Irrtums von der Verantwortlichkeit ist, als welcher auf dem Irrtum von der Freiheit des Willens ruht"* [90].

and brain.

If we wish to understand the concrete importance of our Symposium, we must recall that between a wrong theory and its practical consequences there is a fine line which can be easily crossed. History is full of tragic examples. If we start by rejecting the person on account of a theory based on 'scientific facts', we can end up destroying the radical sense of that term, with the help of a practical ideology of anti-humanism. It is not difficult to find examples such as unchecked genetic manipulation, etc. For this reason, the dialogue between the neurosciences, philosophy, and theology has one point of verification in the capacity to incorporate the ethical criteria necessary for human life in this planet to continue to grow, serving the common good. The dialogue which Heisenberg records between the greatest physicists of the twentieth century can be seen as a beautiful example of a humanist position [36].

Obviously, the dialogue between science, philosophy, and religion serves in the same way to also purify philosophy of its excesses — which have been committed in its theoretical reflection as well as its practical applications — as well as those of religion. Religion is obligated to understand what Benedict XVI has said: "not to act in accordance with reason is contrary to God's nature." Religious belief and Christian faith in particular need a continuous confrontation with truly human reason. And if this affirmation is valid for other realms of theological reflection, it is also valid for the particular mode which we have dealt with in this Symposium, where the advances of the neurosciences provide a continuous stimulus for the progress of philosophical-theological comprehension of man as such, in his consciousness and freedom[51].

### *7.2 The unity of the soul-body duality as an expression of the imago dei*

Christian theology rejects materialistic monism and defends "dual unity." In this way, it situates itself in a humanist perspective[52]. Its point of departure is the revealed doctrine in both the New and the Old Testaments concerning man as the image of God. The biblical tradition also affirms that man has been created in "the image and likeness of God." This is the "base of all Christian anthropology." [91]. Such a conception of the human person does not resolve the soul-body polarity either in the materialist or spiritualist pole, and also avoids the danger of a dualistic interpretation. In the Bible we find a vision of man as a psychosomatic unity (to use scientific terminology), as an animated flesh or as an incarnate soul (to use theological and philosophical terminology). The fundamental implication of Christian revelation is that the human person is what he is through his body, that the human person

[51] *"Es bleibt ein langer und spannender, aber beileibe nicht aussichtsloser Weg, diesen Kulturhistorischen Bestand mit jener Empirie zu vermitteln, die Erfahrungen mit diesem Selbst aus aktutellen neuronalen Prozessen zu erklären versucht"* [92].

[52] It is impossible to present here the biblical, patristic, and magisterial reflection on man as the image of God; see [94, 58, 3].

recognizes that he is inserted into the cosmos and participates in the laws of nature through his sensibility and intelligence. Moreover, thanks to his spirit he can, through his theoretical and ethical-practical faculties, transcend the cosmos and participate in a spiritual dimension that brings him together with other men and women[53].

If we wish to summarize the greatest teaching of the Church on this topic, we can do it with the following affirmations: that man, created as image of God, is a unity constituted by soul and body[54] [95]; that the soul is for itself and essentially the form of the body[55]; that the soul is spiritual[56] and immortal[57].

Theology coined the formulation *anima forma corporis* to express this mysterious dual unity of soul and body that is characteristic of man as a creature of God. It means that the human person is a substantial unity of mind and body. It was Thomas Aquinas who especially explored its contents. The dominican theologian asserts that the soul is the body's unique form. In this way he defends the notion that the human person is not constituted by a mere juxtaposition of two realities complete in and of themselves, which in some way or another are predestined to be united and which unite themselves extrinsically. Instead, soul and body are two principles by virtue of which the human person exists in his original unity[58].

The sense of the unity of the human person starting from the affirmation of the soul as *unica forma corporis* comes with the adoption of the Aristotelian formula: *"anima quodammodo omnia"*[59] in Thomas Aquinas[60]. In this perspective we can also read some other phrases: *"anima enim est in corpore ut continens, et non ut contenta"*[61]; *"non enim anima continetur a corpore, sed potius continet corpus"*[62]. For this reason, for Thomas, the clear result is that, without a dual unity of soul and body, the human being could not be the image of God[63]. As is well known, in the thirteenth century the debate about these questions was often incendiary and Thomas found more than a little opposition to his view[64]. He took as his starting point

[53]The first traditions in Christian theology, between the second and third centuries A.D., proposed different interpretations of revealed truths. Among them, the Asian school stands out (Justin Martyr, Irenaeus, Tertullian) due to its accurate appreciation of the unity and distinction between soul and body. On the other hand, the Alexandrian school (Origen) tended to emphasize the importance of the soul to the detriment of the value of the body, in order to express the *imago Dei*. A few centuries later, Augustine would also reflect in his thought a tension between the originality of biblical data and the influence of Neoplatonic thought [3].

[54]Cf. H. Denzinger, P. Hünermann, *El Magisterio de la Iglesia* (Herder: Barcelona) 22000 (DH) 800; 1440; 3002.

[55]DH 902.

[56]DH 372; 1440; 2812.

[57]DH 1440; 2766.

[58]1SN, d. 5, a. 3; 1ST, q. 75, q. 76 aa. 1-3; cfr. [96]

[59]*De anima*, 8, 1, 431b 21.

[60]*In libro de Anima*, III, l. 13, n. 787; also 1ST, q. 16, a. 3, co, q. 80, 1, co; q. 84, a. 2, ra 2um

[61]1ST, q. 52, a. 1, co

[62]4ST, q. 62, a. 3, ag 3.

[63]*"Anima corpori unita plus assimilatur Deo quam a corpore separata, quia perfectius habet suam naturam"* (QDP, q. 5, a. 10, ra 5um).

[64]A presentation of the debate can be found in [93].

Aristotelian anthropology but he had to reinterpret it. The advantage of Aristotelianism is that it affirms the substantial unity of these two constitutive components of the person, but his thought ran the risk of placing into doubt the immortality of the soul, because if death decomposed the unity of the two principles, neither of the two could survive on their own[65]. The response which Avicenna gave was unsatisfactory, because it conformed itself with defending only the accidental union between soul and body, both being substantially independent. In contrast, Thomas defends the thesis that the soul is the substantial form of man by providing a new understanding of an idea from Pseudo-Dionysus[66]. Thomas affirms that the intellectual soul is the only form of the human composite, and reunites within it all the forms which are necessary for the constitution of man[67]. Recently, Tobias Kläden has compared Thomas's theological theses with the main contemporary ideas concerning the mind-brain problem, arguing that the formula *anima forma corporis* cannot be explained either in a dualistic or a physicalist way. He concludes that Thomas' point of view can serve to advance contemporary debates [7].

Thomas's doctrine is not limited to the use of the terms soul and body according to the Aristotelian notion of hylemorphism. Thomas overcomes this Greek horizon and gives us a fundamental affirmation for anthropology: the irreducibility of the dual unity of mind and body. We should not say, therefore, that man *has* a soul and body but rather that he *is* an inseparable soul and body. Precisely because of this, man can only reach knowledge about himself and reality if he moves starting from a fact which precedes him. Indeed, if man would want to dispense with his sensibility — which is intimately tied with his corporeality — he could not have any experience of reality nor of alterity. The body is thus the first level in which the human person discovers that he can access alterity, as we saw above. For theological reflection this fact is very valuable because it tells us that if the human person examines the soul-body dual unity, he can discover that he is not at the origin of himself. If, on the other hand, one attempts to explain that dual unity in a monistic or dualistic way, one closes himself off from the possibility of accessing his origin and starting point, given that one is censoring the duality by the means with which one has entered into it.

Together with this theory, which underscores the mysterious unity of man within the duality of his dimensions, Church doctrine also maintains that the specific features of the human soul are spirituality and immortality. The spiritual character of the soul should not be explained in a generic sense but rather as the spirit of a body, just as -– on the other hand — "matter" is highly metaphysical and at all "unspiritual" in itself [30]. We can understand that spirituality does not mean only that the soul cannot be reduced to a corporeal dimension

[65] Aristotle resolves the problem with a distinction between man's psyche and nous, and conceived of the latter as separate from the soul, as something general. This separation was unacceptable to Christian thought, because it compromised the individual character of man's nous. Cf. DH 1440.

[66] Cfr. E. H. Weber, *Dialogue et discussion entre S. Bonaventure et S. Thomas d'Aquin à Paris (1252-1273), Vrin, Paris 1974*[93].

[67] The Council of Vienne (1312) would seize on this line of reasoning and teach the doctrine of *"animam intellectivam seu rationalem, ipsum corpus vere per se et essentialiter informantem"*: DH 900.

but rather indicates the presence in man of a determining factor by virtue of which he is distinguished from God but at the same time participates in His spiritual nature. Moreover, given that the soul is the form of the body and is dependent on God, it shows us that the ultimate unity of man and his objective individual consistency are found in his relation to the Creator. As we can see, this conception of spirituality enriches those facts which philosophy of mind has been able to offer us. The soul is the singular and individuating principle of the body, it is not a spiritual universal which every individual participates in (which could be an Aristotelian conclusion) but rather it constitutes every single person and is unrepeatable in each case. This is another way of saying that the soul is the unrepeatable, personal form of eternal relationship with God in every human being[68].

### *7.3 The "dual unity" individual-community as an expression of the imago dei*

We conclude our presentation by returning to the anthropological foundation of trust. The dramatic anthropology proper to the image of God bears the irreducible and relational character of the subject. To illustrate this, Balthasar uses the example of the child and his mother, a relationship which implies a father, a "third party", and is therefore not binary but takes on the shape of a community [98, 99]. The Swiss theologian argues that the dynamism which awakens self-consciousness within the child as a unique spiritual subject occurs through the initiative of the mother, who calls for the child in an embrace of love, communicating her own self in order to inspire a trusting response. Through this encounter the child and his mother discover in an existential way the properties of reality. This is the path through which concrete realities (people and things) are able to manifest the objectivity of reality as a whole. It is not a casual fact that this example of mother and child situates us within the realm of parentage and implies a calling or vocation from another. The awakening and maturing of the knowing subject happens through the intimate relation with an other (first the mother and father, then other human beings) who is constitutive of the being and knowledge of the child. This understanding of the person turns us in favor of a conception of knowledge which does justice to the immediacy of self-consciousness and at the same time to a constitutive alterity. Wittgenstein had consequently insisted on the necessarily communal nature of all language and reason.

In the encounter between the child and his mother, the polarity "individual-community" is manifested existentially, a polarity that is a proper part of all human experience. This key allows us to overcome the contraposition between theories centered unilaterally on the pole of the individual or on the community. What becomes clear in this polarity is the affirmation that the original social nature of the human person does not come at inverse proportion to his individuality, but on the contrary, that it brings about and fulfills this individuality. By virtue of this condition of "dual unity", we can claim a correlation between both poles in which we can defend the value of the subject and its decisive importance (just as the

[68]DH 1440.

main currents of modern philosophy have already done in the past) as well as, at the same time, claim that a relational-social dimension is constitutive of the individual (which has been claimed by many philosophical/theological camps).

If the human person is described in these terms, one can see how human understanding can encompass both strictly individual dimensions — those that come with direct experimentation or evidence reached by the subject — as well as those that come from trust, placed reasonably, on others. It will be this kind of human beings which will be able to respond to the ideal of the scientific research which Baltimore was referring to in his conference at the Whitehead Institute: "The push toward interdisciplinary science is evident everywhere, especially in biology," as he put it.

> *The sequencing of the human genome and the subsequent mining of this information are perfect examples. (...) These large collaborative efforts are generating a tidal wave of data that has required the development of powerful new tools to manage, compile, and manipulate the massive amount of information. As a result, trust is more important than ever.*

He adds, "Of course, along with trust, there must be stringent testing and quality control."[69]

## 8. Conclusion

We started our presentation citing a Nobel laureate in medicine, which opened up our horizons toward trust. We would like to end by citing another Nobel laureate, this one in literature, in order to open our horizon towards eternal meaning and its significance for our personal lives. The poem is titled "Meaning," and the writer is the Polish poet, Czesław Miłosz[70]:

> *When I die, I will see the lining of the world.*
> *The other side, beyond bird, mountain, sunset.*
> *The true meaning, ready to be decoded.*
> *What never added up will add up,*
> *What was incomprehensible will be comprehended.*
> *And if there is no lining to the world?*
> *If a thrush on a branch is not a sign,*
> *But just a thrush on the branch? If night and day*
> *Make no sense following each other?*
> *And on this earth there is nothing except this earth?*
> *Even if that is so, there will remain*
> *A word wakened by lips that perish,*
> *A tireless messenger who runs and runs*
> *Through interstellar fields, through the revolving galaxies,*
> *And calls out, protests, screams.*
>
> [Translated by Santiago Ramos]

[69] Cf. supra nota 1.
[70] CZESŁAW MIŁOSZ, "Meaning" in: New and Collected Poems 1931-2001, Penguin, London 2005. 569.

## 9. Acknowledgements

I am grateful to Marcos Cantos and Eduardo Lostao for their precious help in the preparation of this paper.

## References

1 von Balthasar, H.U. 1976, *Theodramatik II/1* (Johannes Verlag: Einsiedeln).
2 Scola, A. Card. 2010, *"Anima e neuroscienze"*, in: *Christus, Gottes schöpferisches Wort. Festschrift für Christoph Kardinal Schönborn*, eds. G. Augustin et al. (Herder: Freiburg-Basel-Wien), pp. 87-97.
3 Scola, A. Card., G. Marengo and J. Prades 2000, *La Persona Umana. Antropologia Teologica – Manuali di Teologia Cattolica* (Jaca Book: Milano),pp. 141-203.
4 Beorlegui, C. 2010, *Filosofia de la mente* (Panorama), pp. 17-18.
5 Habermas, J. 2006, *Entre Naturalismo y Religión* (Paidos: Barcelona).
6 Wittgenstein, L. 2000, *Sobre la certeza* (Gedisa: Barcelona), par. 115, 150, 378.
7 Hemmerle, K. 2000, *"Verità e Testimonianza"*, in: *Testimonianza e Verità*, ed. P.Ciardella & M. Gronchi (Citta Nuova, Roma), p. 321.
8 González Álvarez, J. 2010, *Breve historia del cérebro* (Critica: Barcelona).
9 Giddens, A. 1990, *The Consequences of Modernity* (Polity Press: Cambridge), pp. 29-36.
10 Clayton, Ph. 2004, *Mind and Emergence. From Quantum to Consciousness* (Oxford Univ. Press: New York), p. 112.
11 Giménez Amaya, J.M. 2010, *"Entender mejor la libertad: un enfoque interdisciplinar entre neurociencia y filosofía"*, in: *Neurofilosofia. Perspectivas contemporáneas*, eds. C. Diosdado, F. Rodriguez Valls and J. Arana (Themata-Plaza y Valdes: Sevilla-Madrid), pp. 177-190.
12 Searle, J. 1986, *Mind, Brains and Science* (Harvard Univ. Press: Boston, Mass.), p. 13.
13 Tallis, R., "What neuroscience cannot tell us about ourselves", The New Atlantis, Fall 2010, pp. 3-25.
14 Chalmers, D. 2002, *"La conciencia"*, in: *Investigación y Ciencia*, Temas 28, pp. 4-6.
15 Millán-Puelles, A. 1997, *El interés por la verdad* (Ediciones Rialp).
16 Habermas, J. 1982, *Conocimiento e interés* (Taurus: Madrid), p. 13.
17 Jaki, S.L. 1983, *Angels, Apes and Men* (Sherwood Sudgen: La Salle, Ill.), p. 99.
18 Popper, K. and J. Eccles 1984, *The Self and its Brain: An argumentation for interactionism* (Routledge), p. 399.
19 Wittgenstein, L. *Tractatus Logico-philosophicus*, 6.52.
20 Steiner, G. 1998, Errata (Siruela:Madrid), pp. 57-67.
21 Giménez Amaya, J.M. And J.I. Murillo 2009, *"Neurociencia y libertad, una aproximacion interdisciplinar"*, in: *Scripta Theologica* 41, pp. 13-46.
22 Giménez Amaya, J.M. And J.I. Murillo 2007, *"Mente y Cerebro en la Neurosciencia Contemporanea. Una Aproximacion interdisciplinar"*, in: *Scripta Theologica* 39, pp. 607-636.

23 Richardson, A. 2003, "The scientific world conception: logical positivism", in: The Cambridge History of Philosophy 1870-1945, vol. 1 (CUP: Cambridge).
24 Kolakowski, L. 1969, *Tratado sobre la mortalidad de la razón* (Monte Avila: Caracas), pp. 298-300.
25 Marcuse, H. 1972, *El Hombre Unidimensional* (Seix Barral: Barcelona), pp. 214-216.
26 Artigas, M. 2007, *Ciencia y Religión. Conceptos Fondamentales* (Eunsa: Pamplona), pp. 67-76.
27 Drees, W.B. 1996, *Science, Religion and Naturalism* (Cambridge U.P.: Cambridge).
28 Haught, J.F. 2006, *Is Nature enough? Meaning and Truth in the Age of Science* (Cambridge U.P.: Cambridge).
29 Bormann, F.J. and B. Irlenborn, eds. 2008, *Religöse Überzeugungen und öffentliche Vernunft. Zur Rolle des Christentums in der pluralistischen Gesellschaft* (Herder: Feriburg-Basel-Wien).
30 Spaemann, R. 2000, *Die Seele: Metapher oder Wirklichkeit? Philosophische Ergruendungen Nikl*, Eds. P. and G. Terizakis (Transcript:Bielfeld), p. 47.
31 Husserl, E. 1982, *Investigaciones Lógicas* (Alianza: Madrid)., pp. 35-188.
32 Merleau-Ponty, M. 1945, *La Phénoménologie de la perception* (Gallimard: Paris) II-III.
33 de la Peña, R. 1988, *Teología de la Creación* (Sal Terrae: Santander), p. 209.
34 Scheler, M. 2005, *Die Stellung des Menschen im Kosmos* (Bouvier Verlag: Bonn).
35 Habermas, J. 2001, *Glauben und Wissen* (Surhkamp: Frankfurt), p. 20.
36 Heisenberg, W. 1972, *Diálogos sobre la física atómica* (BAC: Madrid).
37 Ratzinger, J. Card. 1998, *Glaube zwischen Vernunft und Gefuehl* (Mitteilungen des Übersee-Clubs, Hamburg), pp. 1-21.
38 Prieto Lopez, L. 2008, *El hombre y el animal* (BAC: Madrid).
39 Gehlen, A. 1950, *Der Mensch, seine Natur und sein Stellung in der Welt* (Athenaeum-Verlag: Bonn).
40 Plessner, H. 1975, *Die Stufen des Organischen und der Mensch. Einleitung in die philosophische Anthropologie* (De Gruiter: New York-Berlin).
41 von Balthasar, H.U. 1953, *Der Christ und die Angst* (Johannes Verlag: Einsielden).
42 Gasser, G. and J. Quitterer (Eds.) 2010, *"Die Aktualität des Seelenbegriffs."*, in: *Interdisziplinäre Zugänge* (Schoening – Paderborn).
43 Coady, C.A.J. 2010, *Testimony. A Philosophical Study* (Clarendon Press: Oxford). pp. 54ss.
44 Lackey, J. 2010, *Learning from Words: Testimony as a Source of Knowledge* (OUP: Oxford).
45 Bruaire, C. 1972, *"Témoignage et Raison"*, in: *La Testimonianza*, ed. E. Castelli (Inst. Studi Fil., Roma) p. 146.
46 Klaeden, T. 2011, *"Anima forma Corporis. Zur Aktualitat der nichtdualistischen Sicht des Menschen bei Thomas von Aquin"*, in: *Die Aktualitat des Seelenbegriffs*, ed. Gasser-Quitterer, pp. 253-270.
47 Angelini, G. 2009, *"La Testimonianza"*, in: *La testimonianza cristiana e testimonianza di Gesù alla verità*, ed.G. Angelini & S. Ubbiali (Glossa: Milano), p. 39.

48 Angelini, G. 2009, *"Le Resistenze Antiche: l'antropologia delle facolta"*, in: *La testimonianza cristiana e testimonianza di Gesù alla verità*, ed. G. Angelini & S. Ubbiali (Glossa: Milano), p. 55.
49 Goldberg, S.C. 2011, *Relying on Others: An Essay in Epistemology* (OUP: Oxford).
50 Luckman, T. 2002, *Wissen und Gesellschaft* (UVK: Konstanz).
51 Hierro Pescador, J. 2005, *Filosofía de la Mente y de la Ciencía Cognitiva* (Akal: Madrid).
52 Martínez Freire, P. 2005, *La importancia del conocimiento. Filososfia y ciencias cognitivas* (Thema-Universidad de Malaga: Malaga).
53 Bennett, M.R. & P.M.S. Hacker 2003, *Philosophical Foundations of Neuroscience* (Wiley-Blackwell).
54 Rager, G. 2000, *"Bewusstsein und Person in Wissenschaft und Lebenswelt"*, in: *Bewusstsein und Person*, eds. Rager & Holderegger, pp. 9-26.
55 Crick, F. 1993, *Astonishing Hypothesis: The Scientific Search for the Soul* (Simon & Schuster: London).
56 Penfield, W. 1975, *The Mystery of the Mind: A Critical Study of Consciousness and the Human Brain* (Princeton Univ. Press: Princeton).
57 Candela, J.A., C. Cañón and A. Hortal 1986, *"Monismos, Dualismos y Emergentismos"*, in: *Mente y Cuerpo*, ed. A. Dou (Mensajero: Bilbao) pp. 19-55.
58 de la Peña, R. 1983, *Las Nuevas antropologías: Un reto a la Teología* (Sal Terrae), pp. 219-232.
59 Bunge, M. 1980, *El problema mente-cerebro* (Tecnos: Madrid).
60 Mora, F. 2004, *"Cerebro y Mente"*, in: *Esplendores y Miserias del Cerebro*, ed. F. Mora (Fundación Santander Hispano: Madrid) pp. 203-225.
61 Midgley, M. 1992, *Science as Salvation. A Modern Myth and its Meaning* (Routledge: New York).
62 Eccles, J. 1991, Evolution of the Brain: Creation of the Self (Routledge: London & New York), p. 237.
63 Pankseep, J. 1998, *Affective Neuroscience: the foundation of human and animal emotions* (Oxford Univ. Press: New York).
64 Fodor, J. 1968, *Psychological Explanations* (Random House).
65 Putman, H. 1967, *Psychological Predicates*, in: Art, Mind and Religion, ed. W.H. Captain, (Pittsburgh).
66 Jeeves, M. & Brown, W.S. 2008, *Neuroscience, Psychology and Religion: Illusions, Delusions and Realities about Human Nature* (Templeton Press: Radnor, PA).
67 Laín Entralgo, P. 1991, *Cuerpo y alma* (Espasa-Calpe:Madrid).
68 Varela, F. 2004, *"Cerebro y conciencia"*, in: *Mora, Esplendores y miserias del cerebro* (Fund. Banco Santander), pp. 226-252.
69 Kurthen, M. 2008, *"Neurowissenschaft des Selbst"*, in: *Die Aktualität des Seelenbegriffs*, eds. Gasser & Quitterer, pp. 155-189.
70 Millán-Puelles, A. 1967, *La structura de la subjetividad* (Ediciones Rialp), 184ss.
71 Prades, J. 2008, *"Un testigo eficaz: Benedicto XVI"*, in: *AA.VV. Dios salve la razón* (Encuentro: Madrid), pp. 7-27.

72 Buber, M. 1993, *Yo y tu* (FCE: Madrid).
73 Ricoeur, P. 1995, *Soi meme comme un autre* (Seuil: Paris).
74 Zubiri, X. 2006, *Tres dimensiones del ser humano: individual, social, histórica* (Allianza: Madrid).
75 Damasio, A. 1994, *Decartes's Error: Emotion, Reason and the Human Brain* (Harper Collins: New York).
76 Gazzaniga, M. 2005, *The Ethical Brain* (The Dana Press: New York).
77 Gazzaniga, M. 1998, *The Mind's Past*, (Univ. California Press: Berkeley).
78 Feito Grande, L. 2011, *"Neurociencia de las emociones: claves del comportamiento humano"*, *Di'alogo Filosófico*, 27, 225-241.
79 Sperry, P. 1983, *Science and Moral Priority* (Columbia U.P.: New York).
80 Quitterer, J. 2000, *"Lebensprinzip des Organismus. Ergebnisse der Hirnforschung im Licht von Philosophie und Theologie"*, in: *Herder Korrespondenz* 54, pp. 404-407.
81 Ledoux, J.E. 1996, *The Emotional Brain* (Simon & Schuster: New York).
82 Ledoux, J.E. 1993, "Emotional Network in the Brain", in: Handbook of Emotions, eds. M. Lewis & J.M. Haviland (The Gilford Press: New York).
83 Goleman, D. 1998, *Working with Emmotional Intelligence* (Bloomsbury: London).
84 Goleman, D. 1995, *Emmotional Intelligence* (Bloomsbury: London).
85 Goldman, A. & D. Whitcomb, Eds. 2011, *Social Epistemology: Essential Readings* (OUP: Oxford).
86 Heidegger, M. 1977, *"Sein und Zeit"*, in: *Gesamtausgabe Bd. 2 XIV*, ed. F.-W. Von Hermann.
87 Pérez Ransanz, A.R. 2011, *"El papel de las emociones en la producción de conocimiento"*, *Estudios Filosóficos*, 60, 51-64.
88 Lewis, C.S. 1943, The Abolition of Man (HarperCollins: New York).
89 Honnefelder, L. 2000, *"Der Begriff der Person in der aktuellen ethischen Debatte"*, in: *Bewusstsein und Person*, eds. Rager & Holderegger (Universitaetsverlag Freiburg), pp. 53-69.
90 Nietzsche, F. 1978, *"Menschliches, Allzumenschliches"*, in: *Werke*, ed. von K. Schlechta (Hanser: Muenchen), I, p. 453.
91 John Paul PP. II 1988, *Mulieris Dignitatem* – Apostolic Letter (Libreria Editrice Vaticana).
92 Holderegger, A. 2000, *"Person in der Perspektive von Theologie und Ethik"*, in: *Bewusstsein und Person*, eds. Rager & Holderegger (Universitaetsverlag Feriburg), pp. 86-105.
93 Weber, E.H. 1975, *"Dialogue et discussion entre S. Bonaventure et St. Thomas d'Aquin à Paris"* (1252-1273) (Vrin:Paris).
94 Comisión Teológica Internacional 1998, *"Algunas cuestiones actuales de escatologia"*, in: *Documentos (1969-1996)* (Madrid).
95 Denzinger, H. and P. Hunermann (DH) 2000, *El Magisterio de la Iglesia* (Herder: Barcelona).
96 Bernath, K. 1969, *Anima Forma Corporis. Eine Untersuchung über die ontologischen Grundlangen der Anthropologie des Thomas von Aquin* (H. Bouvier: Bonn).
97 Weber, E.H. 1970, *L'homme en discussion à l'Universitè de Paris en 1270* (Vrin: Paris).

98 von Balthasar, H.U. 1990, *Mein Werk. Durchblicke* (Johannes Verlag: Freiburg), p. 92.
99 von Balthasar, H.U. 1967, *"Der Zugang zur Wirklichkeit Gottes"*, in: *Mysterium Salutis II* (Benziger), pp. 15-45.
100 Barbour, I.G. 2004, *Religión y Ciencia* (Trotta: Madrid), p. 429.
101 Davidson, R.J. 2002, "Seven sins in the study of emotion: correctives from affective neuroscience", Brain and Cognition, 52, 129-137.
102 Milosz, C. 2005, "Meaning", in: New and Collected Poems 1931-2001 (Penguim: London).
103 Polkinghorne, J.C. 1990, "A Revived Natural Theology", in: Science and Religion. One World: Chaning Perspectives on Reality, J.Fennema and I. Paul, eds. (Kluwer: Dordrecht), p. 88.
104 Ricouer, P. 2009, *L'hermeneutique à l'école de la phenomenologie* (Beauchesne: Paris).