



Statement of the Working Group on Neurosciences and the Human Person



The working group met for three days to discuss questions at the boundaries of neurosciences and philosophy, with an emphasis on areas where the scientific approach is making progress, and which lie at the core of what it means to be a human person: the evolution of the human brain, the mechanisms of consciousness, the capacity for evaluation and decision making and self-control, the formation of beliefs in a social group, the sense of self, and the importance of education for human brain development. For each of these topics, we summarize here the essential conclusions and the potential points of convergence between the scientific and philosophical approaches, without denying that many of these points remained heavily debated.

Human brain evolution

Paleontological evidence, exploring the consequences of climate change, nutrition and human migrations, together with genetic evidence, pointing to a limited number of recent mutations unique to the human lineage, are shedding new light on the origins of *Homo sapiens*. Human and non-human primates share brain mechanisms, both at the level of individual circuits and areas and in the manner in which these spatially distributed systems interact and are bound together by synchronized oscillatory mechanisms to form global brain-scale assemblies. The latter mechanisms play a prominent role in processes of computation, arousal, attention, and conscious perception.

The increasing complexity of the human brain led to the emergence of novel cognitive and executive abilities that enabled *Homo sapiens* to engage in cultural evolution. Although rudimentary forms of culture, including intergeneration transmission of ways of cracking nuts, etc,

have been demonstrated in apes, essential steps in this process were the conception of tools, the growing awareness of a finite life span, the development of a symbolic communication system, the transgenerational transmission of acquired knowledge by education, the creation of social belief and value systems, social cooperation, and the concretization of mental representation systems in rituals, artistic endeavors, and social institutions.

This paleontological research, although still developing, represents a major progress over the Medieval vision of the brain, which St Thomas summarizes as follows: 'For man needs the largest brain as compared to the body; both for his greater freedom of action in the interior powers required for the intellectual operations; and in order that the low temperature of the brain may temper the heat of the heart, which has to be considerable in man for him to be able to stand erect' (Thomas Aquinas, *S.Th.* I, 91, 3 ad 1). Already Plato had the intuition that the dimension that makes humans distinct from animals and plants, addressed as mental or spiritual, is a consequence of the evolution of the brain and the ensuing cultural constructs (cf. *Timaeus*, 90 a-b).

Consciousness

Elementary mechanisms of consciousness are now increasingly being analyzed at the brain level. The events that human observers report as conscious differ from those that they cannot consciously report in multiple objective neurophysiological parameters. The sense of ownership of our body and actions, and the feeling of the first-person perspective (or self-perspective) and interoception, can be manipulated in the laboratory, and their neural correlates are being discovered using brain imaging.

The sense of the unity of consciousness, although partially illusory, may be understood as a process of convergence of the various cognitive abilities or "modules" into a single large-scale brain network or "workspace". This integration, made possible by the existence of long- and short-distance anatomical and functional connections linking higher-level brain areas, is thought to enable an internal synthesis of the innate and acquired dispositions of the individual and of his recognition of himself and his body in the world.

The practical implications of these findings are important in medical practice, in order to better understand the loss of consciousness during anesthesia or following brain lesions, to facilitate the detection of residual consciousness in locked-in patients, and to search for mechanisms underlying the disruption of the unity of consciousness in psychiatric illnesses such as schizophrenia.

The discovery that consciousness can be related to specific brain systems should not conceal the fact that most of our brain operates non-consciously. Neuroscience and psychology have discovered that many of the brain processes for knowledge, value, decision, belief formation and social representation reside at a deep and unconscious level that remains unavailable to accurate introspection and conscious recollection. Nevertheless, through self-observation, we develop some degree of explicit self-knowledge, as well as explicit theories on how our mind and the minds of others work (theory of mind).

Values and decisions

How human and non-human primates attribute values and take decisions can also be related to a set of brain areas that provide a prospective evaluation of the consequences of actions at multiple levels. Some of these areas, which can evaluate purely imaginary and novel decisions, are the brain areas most evolved in humans compared to other primates, suggesting an expanded functionality of the valuation and decision system during the evolution of hominins. The capacity for decision making exhibits an inter-individual variability which can be correlated with neurotransmitter concentrations and with functional connectivity between brain regions and networks. At one extreme, addicted individuals show a massive disequilibrium in dopaminergic networks, which biases their entire decision system towards the pursuit of drug taking while undermining their ability for self-control. Understanding of these circuits opens up the possibility of a future treatment by restoring the operation of the valuation system and strengthening that of self-control.

Beliefs and socialization

The formation of belief systems is another cognitive domain of exceptional expansion in the human species. Perceptions and beliefs are thought to jointly arise from hierarchical brain networks that confront internal models of the world with external signals and use the corresponding error signals as a corrective mechanism. Schizophrenia is a mental illness which can be analyzed as the abnormal operation of this error-propagation device.

Although hierarchical perception and belief systems exist in monkeys and apes, they reach their full development in humans in whom an additional level of sharing attention and social information plays an essential part. The circuits of the human social brain, which have only recently started to be explored, may shed some light on how we generate an interpretation of our own self, our behavior, and our sense of responsibility and accountability. The strong feeling of belonging to a social group arises extremely early during development: even infants already express preferences for others who speak the same language. The human sense of belonging to groups generates powerful social tendencies towards in-group cooperation and out-group exclusion.

How altruistic cooperation could have evolved is the subject of several mathematical models. They suggest that cooperation is a genuine force in evolution that may even be necessary for the emergence of any complexity in life, from pluricellular organisms to insect societies and human language. In these models, trust, generosity and forgiveness enter explicitly as evolved traits that stabilize group cooperation. Humans may be distinguished by a specific form of cooperation, 'indirect reciprocity', which relies on language to extend cooperation to new individuals based solely on their social reputation. In this respect, the cognitive sciences strengthen in a new and genuine way the philosophical notions that are at the foundation of ethical and political systems, according to which a human being is essentially a 'social animal' (Aristotle, *Politics*, I, 9, 1253 a 2).

The fundamental importance of education

Even if they are not entirely accurate, human mental representations develop continuously and can be improved by education, an activity that may well be unique to humans. Brain

neuroplasticity is the mechanism by which new memories and learning occur in the brain. In the human brain it allows us to not only transmit tradition and knowledge through education, but also to shape and form personality traits. Education even permits fighting against some of the now maladaptive traits that the brain inherited from its evolution. Even a few weeks of training can modify the brain networks for attention and self-control, thus enhancing willful action over automatic reaction. Thus, any human has in himself an active capability that allows him to progress 'towards himself and his own perfection' (Aristotle, *De Anima*, II, 5, 417 b 3). Hence the importance of educating and 'training' perception, knowledge, reasoning and action, in order to attain truth, good and justice.

Developing a better understanding of how knowledge, action and decision circuits diversify to enlarge the repertoire of our species, as well as studying how the teacher's brain operates to transmit education, are essential goals for future research. The teacher's role in Plato and Thomas Aquinas is to be the instrument that helps their students not only to lead out (*educere*) their own abilities, but to also develop knowledge for themselves. There is a possible point of convergence here with theories of brain development that attribute, even to very young infants, a vast repertoire of knowledge (about objects, space, time, numbers, language...) and the ability to learn by selecting the most pertinent of these internal representations. Aristotle, somewhat similarly, identifies a distinction between potency and act: knowledge pre-exists in the learner in an active way, not passive as it is in general in matter. Otherwise, the human being would not be able to acquire knowledge by himself. Therefore, 'when something pre-exists in the subject in active completed potency, the external agent acts only by helping the internal agent' (Thomas Aquinas, *De Magistro – De Veritate*, 11, 1).

Interdisciplinary convergence and its difficulties

One of the most complex questions in the interdisciplinary approach is to clarify the often different meanings that a word can have when used by the different disciplines. For instance, the analysis of the 'self' is a privileged subject for interactions between neuroscience, psychology and philosophy, reflecting the different traditions and competences of these disciplines. However, they differ with respect to the epistemological status accorded to the object of investigation, the language used for descriptions and the applied methods of research. Kant in this context distinguishes between a determining self (thought) and a determinable self (the thinking subject). The neural sciences have as their object the material aspects of the brain and the cognitive and executive functions that depend on the brain, and aim to explain various aspects of thought on this material and computational basis. Metaphysical anthropology, however, draws attention to aspects of the subject that, it suggests, may not only be thought of without matter but can also be without matter. For example, based on activities such as the perception of time – dissociated from the characteristic succession of physical movement and associated to the mental principles typical of human praxis – and the insatiable thirst for knowledge, life and happiness, realist philosophy (past and present) considers that knowing the truth and tending towards good and justice are specific to human beings. The human person, through knowledge and will, draws on the absolute and does not stop at material realities but aspires to symbolic understanding, science and perfect

knowledge, and desires non-market goods, that is, goods without a price, such as dignity, mutual esteem, and happiness. In the future, examining to what extent these non-material aspirations can be linked to the material reality of the brain does not appear to be an inaccessible goal and, on the contrary, should become an important object of study for cognitive neuroscience.

Science has confirmed the existence of trillions of connections between the billions of neurons and neuronal circuits that make up the human brain, and their ramifications inside the body.

Nevertheless, in general philosophers of the Socratic tradition do not agree that this leads to the conclusion that human intelligence and will are just neural events that happen in the brain. For neuroscientists, the brain integrates all bodily functions. From the point of view of the philosophers at the meeting, this does not mean that it gives the body its ontological vital unity, which is given by the soul: '*vivere viventibus est esse*' (Aristotle, *De Anima*, II, 4, 415 b 12). For Thomas Aquinas (and contemporary thinkers of his school), this emergence or independence in acting reveals the independence of being. The being (*esse, actus essendi*) does not belong to the compound but to the intellectual soul proper (the soul subsists in its *esse*, which it communicates to the body and takes back when the body dies and ceases to 'exist'). The being (to be more precise, *esse as actus essendi*) adheres immediately and thus inseparably to the subsistent form. Consequently, the human soul is thought to be incorruptible and thus immortal, created directly and individually by God.

This philosophical conception, especially the central question of the relationship between the brain and the soul, generated intense debates amongst the scientists and philosophers participating in the working group. It was pointed out by the philosophers that brain functions alone may not be sufficient to serve as a basis for ethical and ontological statements about the status of the human person. Humans with severe impairments of brain functions cannot be denied humanity and dignity. Therefore, although the scientists and the philosophers agreed on the fact that the brain gives vital unity, the philosophers' stance was that the soul is the principle differentiating between living beings and is the unifying essence. Whereas organs, the brain included, and the potencies (i.e. the intellect, the will, the senses) are called secondary principles of unity, of coordination and of operation, the individual is the first principle of action and attribution. In the perspective of the neuroscientists present at the meeting, autonomous action and self-attribution could arise solely from the spontaneous patterns of brain activity that auto-organize to provide internal models and motivations to act, including moral operations (behaviours and emotions). In the perspective of philosophers present at the meeting, autonomous action and auto-organisation is the characteristic of living beings (Aristotle, *De Anima*, 412 a 12), and many of them, like microorganisms and plants, do not have a brain but a substantial principle of unity which is the soul. So the soul is the subject but in an active and coordinative sense in living beings of the different species which becomes – in the human person – a principle and responsible subject capable of reflecting on himself/herself. Both perspectives, however, agree that "the brain acts as the neural central driving force of existence" and that "brain death is the death of the individual", as stated in the Pontifical Academy's Statement 'Why the Concept of Death is Valid as a Definition of Brain Death' (2008).

Conclusions

In conclusion, the current knowledge of the organization of the human brain and how it gives rise to mental states already provides an important contribution to the issue of what the human person is. Yet like any scientific enterprise, the answers that it provides remain limited. Scientists and philosophers need to search for a better language that may bridge the gaps between the disciplines and levels of analysis. This includes the language of values, responsibility, dignity and justice. The reconstructions of the concepts of consciousness and self-consciousness, mind and soul, form and information, may help to bring together the natural sciences, the social sciences and the humanities.

Thanks to the discovery of the centrality of the brain, made by the neurosciences, we now have a new starting point for our recognition of the status of the human being. Today, we can be both actors and spectators of our own actions and of ourselves – the first-person perspective of the subjective self is complemented by the third-person perspective of neuroscience. Only a human being is capable of creating such circularity by observing the functioning of his brain from the outside with increasingly powerful instruments, while also interpreting these data from inside, based on conscious self-reflection. The consequences of this dual approach are only beginning to be explored.

In addition to contributing to this conceptual search, cognitive neuroscientists also have an important present responsibility with respect to the many challenges raised by the contemporary world. New interfaces will soon link the human brain to computers and robots, alleviating paralysis but also raising difficult ethical issues. The legal system may benefit from, but also be deeply challenged by, our improved understanding of conscious and non-conscious determinants of human behavior. Many existing human institutions, such as the prison system, may ultimately require extensive reconsideration in light of our growing understanding of the human brain and the possibility of changing and educating it. Prison (deprivation of freedom to move) should never be just a punitive institution but also, and above all else, it should protect society against dangerous individuals, act as a deterrent, and be corrective and educational for those who are imprisoned.