Antonio R. Damasio

The Fabric of the Mind:

A Neurobiological Perspective

Fernando H. Lopes da Silva

From Neuronal Networks to Consciousness

A Comment

COLOFON

uitgave
Nederlandse Organisatie voor Wetenschappelijk Onderzoek
Postbus 93138
2509 AC Den Haag
070 344 06 40
e-mail: news@nwo.nl

NRC Handelsblad
Rotterdam

samenstelling en vormgeving
Afdeling Voorlichting en Communicatie NWO

basislayout
Kees Endenburg
Den Haag

druk
Van Marken Delft Drukkers
Delft

januari 2000

NWO/Huygens Lectures

The Netherlands Organization for Scientific Research NWO and the newspaper *NRC Handelsblad* have organised NWO/Huygens Lectures since 1992. The NWO/Huygens Lectures are a series of lectures, providing for the public at large a unique opportunity to listen to and meet famous scientists. The city council of The Hague acts as a host.

For the 1999 lecture NWO invited the American neuroscientist Antonio R. Damasio, Van Allen Professor and Head of the Department of Neurology of the University of Iowa and Adjunct Professor at the Salk Institute for Biological Studies in La Jolla, California (US). Professor Damasio has made an outstanding contribution to the understanding of cognitive processes in terms of brain functions, through a remarkable series of studies with neurological patients affected by brain damage, making use of advanced brain imaging techniques. These studies led him to develop a new way of thinking about how emotions and feelings interact with rational behaviour.

The successful NWO/Huygens Lectures were organised for the eighth time this year. Predecessors of Professor Damasio were in their field equally famous. For the 1998 lecture NWO invited the British computing scientist Professor Tony Hoare of the Oxford University Computing Laboratory, a specialist in the search for bug-free software. The Dutch Professor Jan Bergstra (University of Amsterdam and Utrecht University) acted as co-lecturer.

The series started in 1992 with a lecture of Professor Nathan Glazer (Harvard University) about ethnic minorities. He was followed in 1993 by Professor Bert Bolin, president of the Intergovernmental Panel on Climate Change. The 1994 lecture, about soft materials, was given by Nobel Prize winner in physics Professor Pierre-Gilles de Gennes, from the Collège the France. In 1995 Professor Jonathan D. Spence (Yale University) talked about Chinese and Euro-American relationships. In 1996, the winner of the 1997 Nobel Prize in Medicine, Professor Stanley B. Prusiner (University of California) talked about his prion concept on the spongiform brain diseases in animals and humans like scrapie, BSE and Creutzfeldt-Jakob. Finally, in 1997 the American paleo-ecologist Professor Paul Colinvaux of the Smithsonian Tropical Research Institute in Panama lectured about the ice-age Amazon and the problem of diversity.

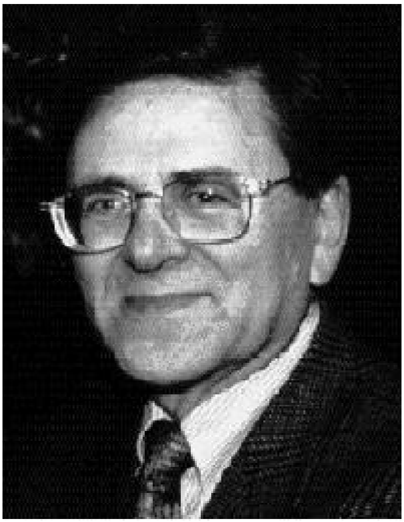
Professor Antonio R. Damasio is Van Allen Distinguished Professor and Head of the Department of Neurology at the University of Iowa and Adjunct Professor at The Salk Institute in La Jolla, California (US). Damasio's work has focused on elucidating critical problems in the fundamental neuroscience of mind and behaviour, at the level of large-scale systems in humans, although his investigations have also encompassed parkinsonism, and Alzheimer's disease. His contributions have had a major influence on our understanding of the neural basis of decision-making, emotion, language, and memory. In collaboration with Hanna Damasio, a distinguished neurologist who is independently recognised for her achievements in neuroimaging and neuroanatomy, Damasio moved lesion studies away from clinical descriptions and placed them at the service of hypothesis-driven research. The laboratories that he and Hanna Damasio created at the University of Iowa are a leading centre for the investigation of cognition using both the lesion method and functional imaging.

Damasio is a member of the National Academy of Sciences' Institute of Medicine; a fellow of the American Academy of Arts and Sciences; a member of the Neurosciences Research Program; a member of the National Advisory Council on Neurological Diseases and Stroke; a fellow of the American Academy of Neurology; a member of the European Academy of Sciences and Arts and of the Royal Academy of Medicine in Belgium; a member of the American Neurological Association, and of the Association of American Physicians, and a board member of leading neuroscience journals. He is a Past President of the Academy of Aphasia and of the Behavioral Neurology Society. Damasio's distinguished lectureships include the Tanner Lecture (Michigan), the Wilson Lecture (Wellesley), the Steubenbord Lectures (Cornell University), the Public Lecture at the Society for Neuroscience, the Aird Lectures (University of California, San Francisco), the Nobel Conference, the Karolinska Research Lecture at the Nobel Forum, and the Presidential Lecture at The University of Iowa. Since 1981 he has delivered an annual series of lectures on the neurology of behaviour at Harvard Medical School. Among others, he has received the William Beaumont Prize form the American Medical Association (1990); the Golden Brain Award (1995); the Ipsen Prize (1997); and the Kappers Medal of Neuroscience (1999). In 1992 he and his wife shared the Pessoa Prize.

Antonio Damasio's book *Descartes' Error: Emotion, Reason and the Human Brain* (Putnam 1994) has been published in over twenty countries. His new book *The Feeling of What Happens: Body, Emotion and the Making of Consciousness* is published by Harcourt Brace.

Antonio Damasio was born in Portugal. He received both his MD and his doctorate from the University of Lisbon, and began his research in cognitive neuroscience with the late Norman Geschwind.

Professor Fernando H. Lopes da Silva received his medical degree from the University of Lisbon in 1959. In 1970, he received his Ph.D. from the Utrecht University with his thesis on System Analysis of Visual Evoked Potentials. In 1973, Lopes da Silva became Head of the Brain Research Department of the Institute of Medical Physics (TNO) in Utrecht. He taught neurophysiology at the Twente University from 1975 to 1985. In 1980, he was appointed full professor of General Physiology at the Faculty of Sciences of the University of Amsterdam, and after the creation of the Institute of Neurobiology at the same University, he became Director of this Institute and member of the Board of Directors of the Graduate School Neurosciences Amsterdam. In 1995, Lopes da Silva was appointed Scientific Director of the Institute for Epilepsy Meer en Bosch in Heemstede.



Lopes da Silva's research interests are mainly the study of brain physiology, including the biophysical aspects of brain electrical/magnetic activity and the functional organisation of neuronal networks. He has a special interest for the phenomena of synaptic plasticity that lead to epileptogenesis and for the analysis of neuronal networks involved in memory processes, particularly of the limbic cortex.

Since 1985 professor Lopes da Silva is a member of the Netherlands Royal Academy of Arts and Sciences. In 1997, he received the Doctor Honoris Causa (Medicine) of the University of Lisbon. He is Honorary Member of the Dutch Society of Clinical Neurophysiology and of the British Society for Clinical Neurophysiology. Most recently he has been awarded the 1999 Herbert H. Jasper Award, presented by the American Clinical Neurophysiology Society for his "lifetime of outstanding contributions to the field of clinical neurophysiology".

Antonio R. Damasio

The Fabric of the Mind:

A Neurobiological Perspective

Abstract It is generally agreed that the set of processes we call mind emerges from the brain, but the question of how that emergence occurs has not been answered. Three centuries ago – just a few blocks from The Hague's New Church where I am delivering this lecture – Spinoza, one of the most remarkable thinkers in human history who was both a contemporary and correspondent of Christiaan Huygens, was preoccupied with this same question. Naturally, Spinoza did not and could not produce a satisfactory answer, but he did venture that mind, on the one hand, and brain/body, on the other, were not separate entities. They were two aspects, two views as it were, of the same essential thing. Spinoza's revolutionary proposal was radically different from those of his contemporaries and, in effect, from those of many generations of philosophical successors up to our own time. His proposal contained the beginnings of an approach into the problem that some of us, in neuroscience or philosophy, now recognize as critical to finding a proper solution.

Spinoza notwithstanding, the question of how the brain/body generates the conscious mind has lingered, off and on, and has come to our time as a major scientific mystery and philosophical puzzle. Lately this question has even become a preoccupation not only for neuroscientists, cognitive scientists, and philosophers, but also for virtually anyone else who wonders about how living matter gives rise to the conscious mind. For some there is little doubt that a proper answer will be found; for others no answer is likely, in principle. In this lecture I argue that a substantial answer to this stirring question will be found, perhaps in the not-too-distant future, and I explain my reasons for believing so.

The timing of a question Let me begin by considering why the question of the origins of the conscious mind has come to center stage at this point in history and not earlier. One likely reason is that biology has had a remarkable success with its discovery of the hidden secrets of life, and that the neurobiological basis of the conscious mind – the modern version of the classic mind-body problem – has virtually become the residual challenge. The relentless and exponential rise of new knowledge has created the vertiginous feeling that no problem can resist the assault of science, if only the theory is right and the techniques are powerful enough. It just seems reasonable to expect that the conscious mind problem will be solved now. Behind that all-conquering giddy feeling, however, lie some notable hurdles that often go unrecognized in the midst of so much enthusiasm.

Acknowledging the hurdles In no particular order of importance, here is a short list of the main hurdles faced by a search for the biological basis for the conscious mind.

The first-person perspective and the problems it poses The perspective we must adopt when we study the conscious mind in relation to the brain, raises some important issues. Both the body and the brain – anyone's body, anyone's brain – are observable by third persons, while the mind is only "observable" by the first-person who owns it. Multiple individuals confronted with the same body or brain can make precisely the same verifiable observations of that body or brain. Yet no comparable direct "third-person" observation is possible for the mind. The mind is a private, hidden, internal, entirely subjective entity. The body and its brain are public, exposed, external, and entirely objective entities. How does the dependence of a first-person mind on a third-person body come to occur? Naysayers such as the philosophers Colin McGuinn and David Chalmers argue that, as neuroscience now stands, once we put together all the images from modern brain imaging scans and from the neurohysiological patterns of activity in brain neurons, all we have discovered and described are the *correlates* of mental states rather than anything that resembles a mental state (McGinn, 1991; Chalmers, 1996). At the conclusion of our observations of living matter we find not the mind but, quite simply, detailed living matter. This argument is fallacious, as I will explain in a moment, but it does silence many of the investigators of the conscious mind.

The intentionality hurdle As a consequence of the first-person perspective problem, the naysayers tell us that we have no hope of understanding how living matter generates the sense of self that hallmarks a conscious mind – the sense that the images in my mind or yours are mine or yours, and are formed in my perspective or yours. It is often stated that the problem is hopeless and that we cannot even explain why the mind is *about* something, in other words, why mental states represent external objects or represent feelings inside an organism. The "about-something-phenomenon" is the property of mind that philosophers have identified by the misleading term "intentionality". The argument that intentionality is an insurmountable hurdle is also false, as we shall see.

The mind is too limited to understand itself Naysayers deliver their last blow when they remind us that the very fact that we can ask the question of how a conscious mind arises in a brain, depends on the existence of that same conscious mind. The investigation of the question is to be conducted with the very same instrument that is being investigated, a situation that makes both the definition and the approach of the problem too complicated for comfort. Because of this conflict between observer and observed, the human intellect my be defeated by the task of understanding how the mind emerges from the brain. I agree that there is a real conflict here but the notion that it cannot be overcome is false.

In conclusion, the apparent uniqueness of the problem of elucidating the basis for the conscious mind and the difficulties that surround its approach are not just daunting for those who are committed to finding a possible solution: they are positively forbidding for those who intuitively never believed a solution could be found. The question of how the conscious mind arises in the brain has finally come to center stage, but there is a considerable resistance to the idea that an answer is possible, let alone likely. There is nothing more familiar than the mind and yet, when it comes to inquiring about its sources and mechanisms, the problem becomes one vast region of strangeness.

Assessing the hurdles Dealing effectively with the conscious mind problem requires a position regarding the difficulties I outlined. I will address first the difficulties that pertain to the circumstances in which the conscious mind is being investigated, for example, the notion that the current understanding of living matter has reached its limits without succeeding in identifying the "substance of mind." This notion is simply incorrect. The current description of neurobiological phenomena is most incomplete. For example, we have yet to unravel numerous details regarding the function of neurons and circuits at the level of molecules. We have yet to understand the behavior of populations of neurons within even small regions of the cerebral cortex, and our understanding of those population behaviors at the level of large-scale systems formed by several brain regions is no less incomplete. We are just beginning to be able to address the fact that pathway interactions among several non-contiguous brain regions probably generate complex biological states that are in no way a mere sum of the local contributions of those several regions. The recent work of Giulio Tononi and Gerald Edelman (Tononi and Edelman, 1998) is an example of these later developments.

It is not reasonable to declare the conscious mind problem insoluble because we have studied the brain to its ultimate border and have not found the mind while doing so. In fact it is not even true that our understanding of matter, in general, let alone our understanding of matter as it relates to biological events is in any way complete. For example, it is conceivable that the comprehensive account of how the brain generates the integrated sensory images that hallmark the mind might gain from a greater understanding of quantum physics. In fact, at the finest level of description, the rapid construction and superposition of sensory images might require explanation at the quantum level. I hasten to add that the acknowledgment of a possible role for quantum physics in the elucidation of the mind, an idea recently identified with the physicist Roger Penrose, should not be confused with an agreement with his specific proposals in this area (Penrose, 1994). For example, Penrose and his colleagues believe that consciousness is based on quantum level phenomena occurring in the subcomponents of neurons known as microtubules, while I do not see the quantum level of operations as necessary at all to explain the defining element of the conscious mind problem which is that pertaining to the self.

In conclusion, I believe we should say that we have not elucidated the full biology of mental phenomena because we have not sufficiently elucidated either neurobiology or its related physics. Perhaps the main reason behind the apparent strangeness of the conscious mind problem has to do with ignorance, plain and simple. Ignorance limits the imagination and seduces otherwise intelligent thinkers into believing, of all things, in magic. Before we knew, based on experimental evidence, the secrets of something as simple as the circulation of blood, there were many bizarre accounts of the phenomenon and much despair. The problem of the conscious mind is far worse because its nature is more complex and the requisite evidence more difficult to obtain.

There is a large obvious gap, no less dramatic than a geological fault, between mental states, on the one hand, and physical/biological phenomena, on the other. The gap comes from the remarkable mismatch of two bodies of knowledge: the fairly well described complexity of "mind processes", and the very defectively described biological processes that correspond to them, that, in fact, *must be* them.

Let me immediately add, lest I may be misunderstood, that I am not *substituting* biological descriptions for mind states, and therefore rejecting the existence of mind. I am affirming the existence of mind processes but suggesting that their nature, their substance, is finely biological. In other words, I am saying that mind as we experience it, is an aspect, a view as it were, of ex-tremely complex processes of living matter, in short the Spinozian proposal to which I alluded in the beginning of the lecture.

In short, there is a mismatch between the form and wealth of phenomena available directly and automatically in the mind processes, and the form and wealth of phenomena in the *current* neuroscientific description of neurobio-logical processes. There is a gap between the mental specification we have achieved through centuries of introspection and the efforts of cognitive science, and the neural specification we have achieved through the efforts of neuroscience. There is no reason to believe, however, that the gap cannot be filled as further research in neurobiology continues.

In principle, nothing indicates that the “mental” and the “neural” must be irrevocably separate.

As for the idea that the real conflict between observer and observed makes the human intellect unfit to study itself, it is important to point out that brain and mind are organized in levels, and the highest of those levels create instruments which permit the observation of the others – reasoning strategies, language, logic, systematic methods of observation and analysis of phenomena. Naturally, this conflict does place a limit on our ability to observe our whole nature, but that is quite different from declaring that we cannot understand any part of it and thus admitting defeat before we even try.

Some possible solutions The solution I have proposed for the problem of the conscious mind calls for separating it into two parts (see Damasio, 1999). The first is the problem of how we generate what I call a “movie-in-the-brain”, an integrated composite of images in diverse sensory modalities – visual, auditory, tactile, olfactory, and so on. The second is the problem of “self”, which pertains to how we generate automatically the sense that we own the movie-in-the-brain. The two problems are related, but they require individual solutions and it is helpful to separate them in terms of research strategy.

A neurobiology for the movie-in-the-brain It is fair to say that for most of the history of neuroscience, we have been attempting to solve the “movie-in-the-brain” problem, a metaphor for the integrated and unified composite of diverse sensory images that constitutes the multimedia mind show. When Paul Broca and Carl Wernicke first suggested that different regions of the brain were involved in processing different aspects of language, a century-and-a-half ago, we had begun mapping brain regions involved in constructing the movie-in-the-brain. In recent years, after the development of more sophisticated tools, the results have become truly impressive. It is possible to record the activity of single neurons or groups of neurons directly, and relate that activity to say, the perception of a color or a curved line or the notion of a point in space. Using PET scans or fMR scans we can determine how the different brain regions of a normal living person are engaged by a mental state – for instance, finding the word specific to a certain object or to a specific person (see for example, H. Damasio et al, 1996). No less importantly, we can investigate how molecules within microscopic neuron circuits participate in such functions, and identify the genes whose activity is necessary for the particular molecules to be formed (Kandel et al., 1999).

Examples of progress in the understanding of the relation between mind and brain abound and have been driving by the availability of new techniques at both ends of the neuroscience spectrum: molecular neurobiology and the combination of functional neuroimaging technologies with cognitive experimentation. These techniques have joined the classical approaches of neuroanatomy, neurophysiology, and neurochemistry and pharmacology, and have helped produce a multidimensional view of the brain engaged in mindful activity.

Examples of progress have not stopped accumulating in the decades since David Hubel and Torsten Wiesel demonstrated in painstaking neurophysiological and neuroanatomical studies that neurons in the primary visual cortex were selectively tuned to respond to edges oriented in varied angles. Their findings gave us a window into the understanding of how brain circuits can represent the shape of an object. Later David Hubel and Margaret Livingstone would show that other neurons in the primary visual cortex respond selectively to color but not to the building blocks of a shape (see Hubel, 1987, for review). In the meantime Semir Zeki showed that brain regions located downstream from the primary visual cortex were relatively specialized for subsequent processing of color or of movement (Zeki, 1993). Interestingly, clinical neurological studies had already shown, in living neurological patients, that damage to distinct regions of the visual cortices would lead, for instance, to a loss of the ability to perceive color while preserving the ability to perceive shape and movement (see Damasio, Tranel and Damasio, 1990 for review). The new neurophysiological and neuroanatomical data from animal studies added microscopic detail to those founding observations. As studies have progressed on both experimental animals as well as neurological patients and normal humans, knowledge of how the brain constructs a visual representation of the objects that impress the retina has revealed a demonstrable correspondence between the structure of a particular object, as presented to the eyes, and the pattern of neuron activity generated by that presentation within the visual cortex of the organism confronted by the object (Tootel, et al., 1988).

Another case of remarkable progress is the understanding of the mechanisms of learning and memory. We have discovered that the brain uses one system to learn facts, pertaining, for instance, to persons, places, and events, and another system to learn skills, for instance, playing a musical instrument or riding a bike. The hippocampus is a key component in the system that serves the learning of facts while the basal ganglia and the cerebellum are critical for the learning of skills (Damasio, Tranel and Damasio, 1989; Tranel et al., 1996; Lopes da Silva, 1996; Lopes da Silva, et al., 1990). We have also discovered that the structures that are so indispensable for learning facts

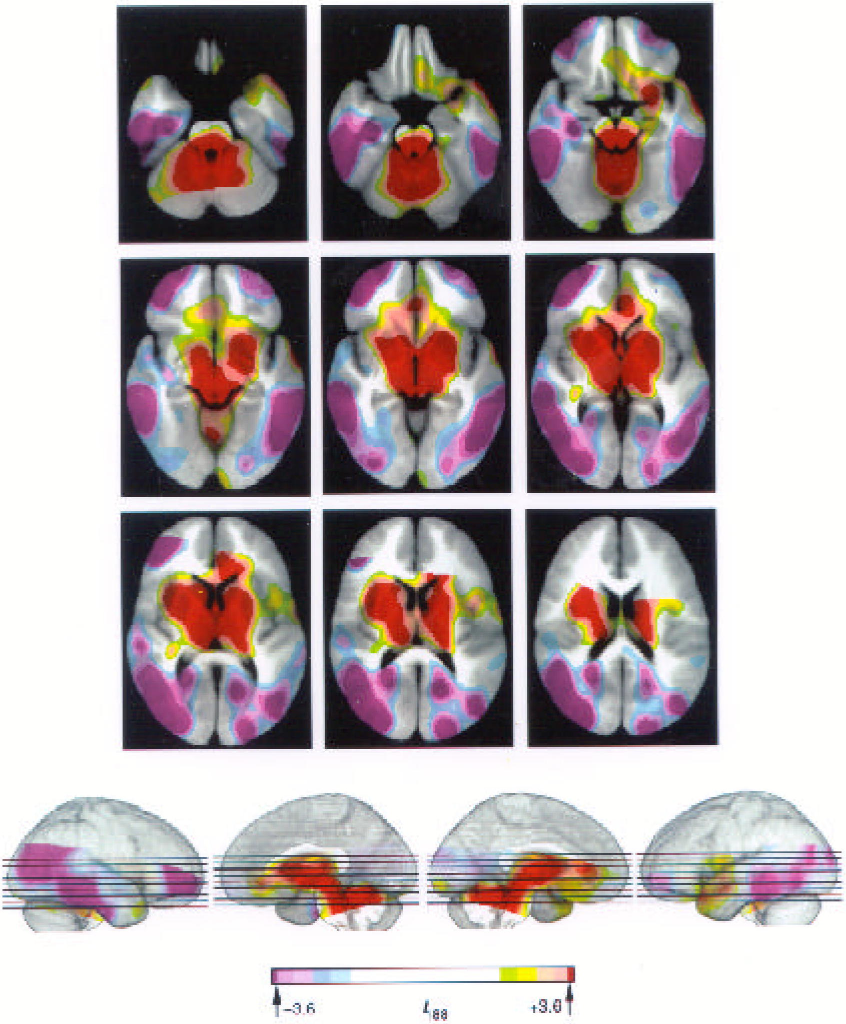


Figure 1. Areas of the human brain activated (red) or deactivated (blue/purple) during the induction of primary emotions and the experience of feelings. The study was performed in normal human volunteers using position emission tomography (PET). The emotions investigated in the study were sadness, happiness, anger and fear.

are not used to store the long-term memories of those facts. It turns out that memories are actually laid down in brain systems made up of many components and that those components are located importantly though not exclusively in the vast brain expanses known as the cerebral cortices (Damasio 1989; Damasio and Damasio, 1994).

We have also discovered that in order for learning to occur, that is, in order for a fact that is now being held in short-term memory to be consolidated in long-term memory, it is not sufficient to have properly working hippocampi and cerebral cortices. It is indispensable that, at the level of neurons and molecules, certain processes take place such that the neural circuits retain the "impressions" caused by the newly processed facts. It has long been known that the retention depends on strengthening or weakening the contacts between neurons – which are known as synapses – but a new and exciting finding is that etching the impression requires the synthesis of new proteins, and that, in turn, requires the engagement of certain genes within the neurons that support the consolidated memory (Kandel et al., 1999).

One topic where recent progress is especially notable is emotion. Emotion was largely neglected by neuroscience during most of the twentieth century, but it is now the focus of intense scrutiny, and not a moment too soon considering its importance in human lives (Damasio, 1994). The neurobiological underpinnings of the emotions have begun to be elucidated and it has become clear that the brain handles different emotions with the help of different components. For instance, from both animal and human studies, it is apparent that the amygdala, a collection of neuron nuclei in the depth of each temporal lobe, is a critical component in the brain system that processes fear (LeDoux, 1995; Adolphs et al., 1998; Adolphs et al., 1997; Adolphs et al., 1994). Surprisingly, not only is the amygdala necessary for us to learn that a certain stimulus can cause fear, but it is also important for us to experience fear in relation to such a stimulus, and even for us to recognize, from a facial expression, that someone else is experiencing fear. Such discoveries, along with the subsequent exploration of other components of the systems related to the processing of fear, are essential for the understanding of a phenomenon such as anxiety which figures prominently in many psychiatric diseases. Moreover, the processing of fear is also a key to a variety of social behaviors, for instance, the judgment of the degree to which another person is trustworthy (Adolphs et al., 1998). Understanding the neurobiology of such systems will play a role in the elucidation of the neurobiology of social navigation.

Recent studies show that fear is not alone in depending on a specific neural system. For instance, sadness recruits the participation of the ventral

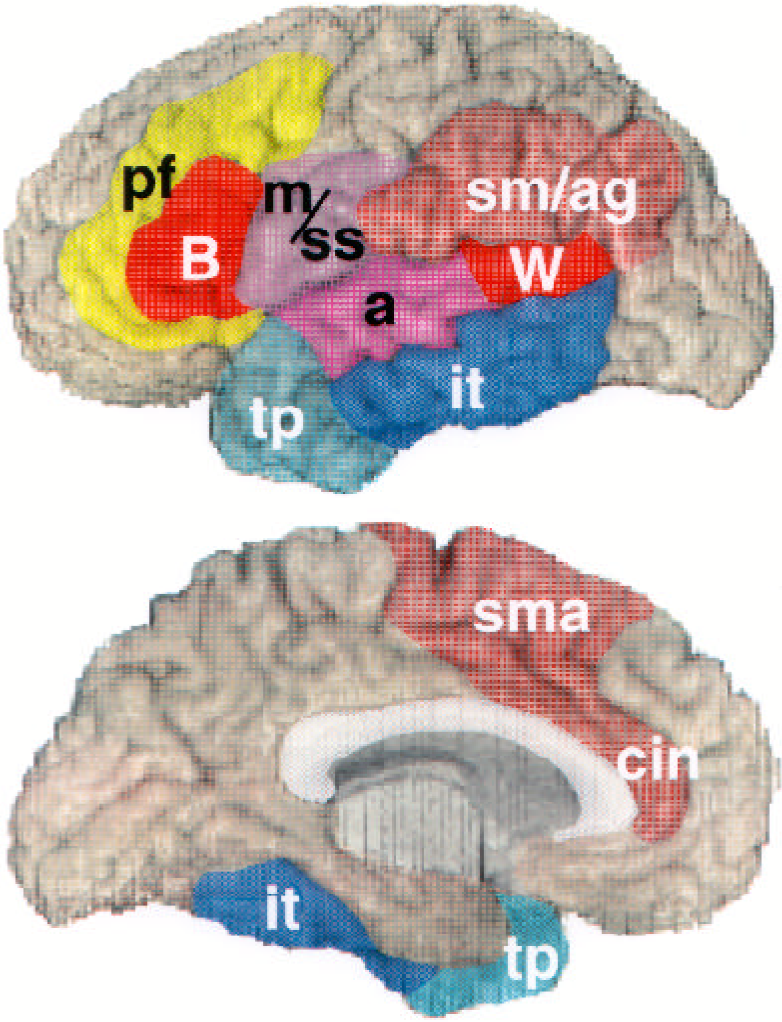


Figure 2. The critical language-related areas of the human brain are located in the lateral (top) and medial (bottom) surfaces of the left hemisphere. B = Broca; W = Wernicke; a = audtiory cortex; tp = temportal pole; it = inferotemporal region; m/ss-motor and somatosensory cortex; pf = prefrontal cortex; sma = supplementary motor area; cin = cingulate cortex. These cortical areas are not neo-phrenological centers but rather key regions of the cerebral cortex. They are subcomponents of several language-related systems.

and medial prefrontal cortex and of the hypothalamus to an extent different from that of other emotions (Damasio et al., 1998). The discovery offers important clues in the investigation of diseases such as depression or mania. For example, in patients with depression and mania there is a malfunction of the sector of ventromedial prefrontal cortex just below the corpus callosum. Remarkably, most of the brain components that in one combination or another form the systems that support each particular emotion, are involved with both the regulation and representation of internal body states – this is true of the hypothalamus, of sectors of the upper brain stem, of basal forebrain, and of somatosensory cortices – underscoring the degree to which emotion and feeling are inextricably interwoven with the mechanisms that ensure the maintenance of life. This finding offers another clue regarding the possible mechanisms of mood disorders and the design of possible interventions.

Yet another example of remarkable progress concerns our understanding of how language is processed in the brain. For over a century we have known that two areas in the human left cerebral hemisphere, Broca's and Wernicke's areas, are involved in both the understanding and the production of language. But from new studies performed with the help of neuroimaging techniques, it is now apparent that numerous other areas of the human brain make up the systems that are involved in producing, quite harmoniously, different aspects of language – the sound structure of words, the grammatical structure of sentences, and the selection of the words with which we name a person, an object, or an action. Over a brief period of time, the map of the language-related brain has gone from being made up of two quasi-phrenological centers, to encompassing a team of cooperating brain regions and systems (Damasio, 1992; Damasio and Damasio, 1992; H. Damasio, 2000).

In conclusion, we can make the following statements. First, it is possible to identify distinct brain regions which work in concert to produce a given mental function. The notion of working in concert needs to be especially emphasized. Throughout this brief review I have been discussing "systems" made up of several brain regions rather than "single" brain regions or areas. There is no example of a complex mental function that can be attributed to – performed by – a single area, in the manner suggested by phrenology a couple of centuries ago. On the other hand it should be noted that the large-scale systems identified by modern lesion studies or functional imaging studies are quite specific in the overall contribution they make to mental processes.

Second, there is a close correspondence between the appearance of a mental state or behavior, on the one hand, and the state of activity of a select number of brain regions. The correspondence can be established both at the level of macroscopic systems: at the level of macroscopically identifiable regions – say, within the visual or auditory cortices, or in emotion-related nuclei; and at the level of the microscopic neuron circuits. But there should be little doubt that the current successes, impressive as they may be, are just a beginning. Techniques are improving both in terms of our ability to analyze neural function at molecular level, and to analyze the high complexity of large-scale phenomena arising from the whole brain. Our ability to establish ever finer correspondences between mental states and brain states will increase correspondingly. The fine grain of physical structures and biological activities which constitute the "movie-in-the-brain" are likely to be gradually revealed as evidence accumulates.

The problem of self The sort of progress outlined so far has convinced many doubters that the biological substrate for the movie-in-the-brain can be discovered. Yet the naysayers have remained unconvinced that the second part of the conscious mind problem – the sense of self – can be elucidated. I grant then that solving this part of the problem is anything but easy but I claim that there are possible solutions and in this lecture I outline a hypothesis that is now being tested.

Let me begin by discussing some fundamental assumptions behind the hypothesis. First, unlike cells in the liver or lung, which perform their assigned biological business but do not *represent* any other cells or business, neurons in the brain *represent* "objects" or "events" occurring elsewhere in the organism. Neurons are *commanded* by biological design to be *about* other cells and other actions. I have said elsewhere that they are born cartographers of the geography of an organism and of the events that take place within that geography. The mystery of the "intentional" mind relative to the objects external to an individual, is hardly mysterious when one realizes that the brain is in the business of directly representing the organism, and of indirectly representing whatever the organism interacts with. There is no reason for the philosophical despair that surrounds the so-called "intentionality hurdle".

Second, the brain's spontaneous "intentionality" hinges on another important aspect of brain design. The brain possesses devices that are aimed at managing the life of the organism in such a way that the internal chemical balances indispensable for survival can be maintained continuously. The brain devices that regulate the life state also represent, of necessity, the continuously changing states of the organism as they occur incessantly. These devices are not hypothetical. They exist and are located in the brainstem nuclei and the hypothalamus and the basal forebrain. The brain has a

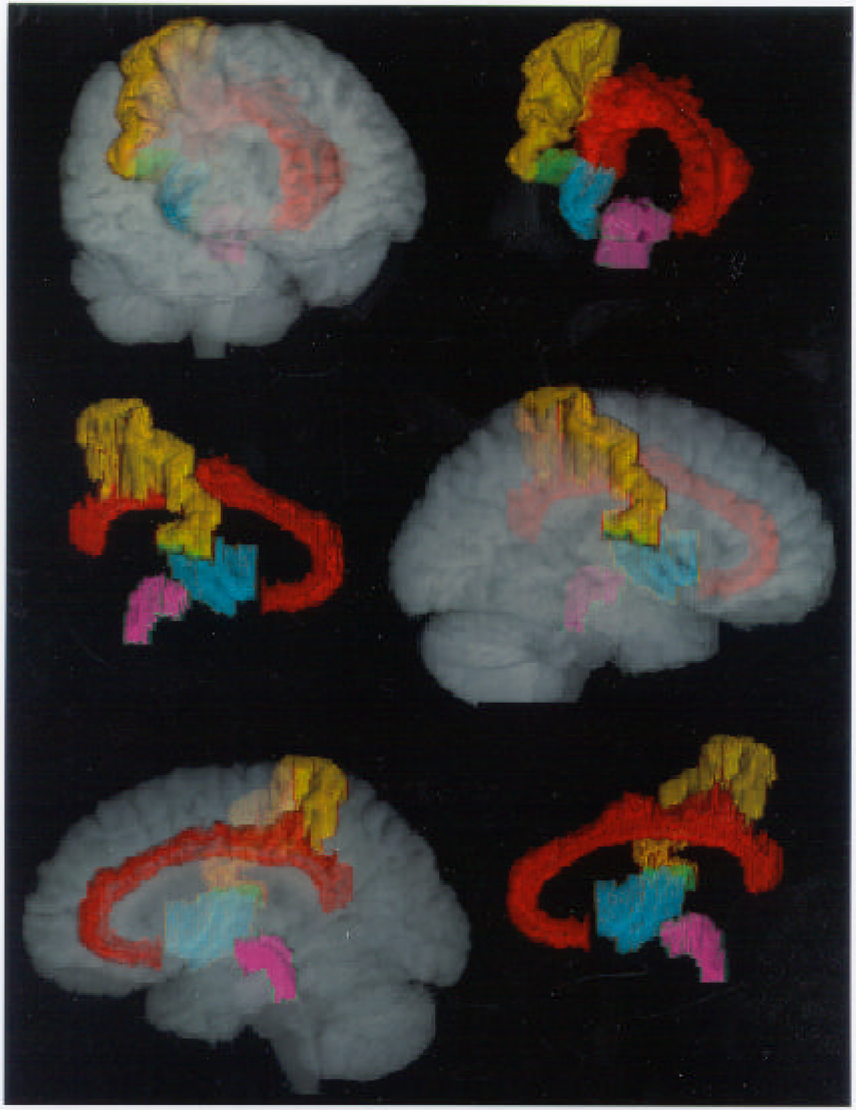


Figure 3. Visualization of brain regions critically involved in the making of core consciousness and the generation of the sense of self (areas in color). Most of these regions are evolutionarily old, and are located near the brain's midline.

natural means to represent the entire anatomy and the current functional state of the whole living organism. I have argued, both in *Descartes' Error* and more recently in *The Feeling of What Happens*, that the biological foundation for the sense of self can be found in these brain devices which represent moment by moment the continuity of the same individual organism.

We should ask, at this point, if it is possible to move from such a biological conception of self to the sense of ownership of one's thoughts which we directly obtain in our mind, without invoking the infamous homunculus who would interpret the situation.

In the simplest of outlines, my hypothesis suggests that *the brain uses mapping structures informed by both maps of the organism and maps of objects, to create a fresh second-order representation which indicates that the organism, as represented in the brain, is involved in interacting with an object. The second-order representation occurs in neural structures such as the thalamus and the cingulate cortices.* This newly constructed knowledge is core consciousness with its core self. The new knowledge *introduced* in the ongoing mental process adds important new information. The new knowledge *presents* within the mental process the information that the organism is the owner of the ongoing mental process. The new knowledge provides spontaneously an answer to a question that the organism never posed, namely, who owns these thoughts? In answering that unasked question the new knowledge creates the sense of a first-person, a sense of self in the act of knowing, thus providing a solution for the alleged impossibility of a biological substrate for subjectivity.

Returning to the metaphor of the mental process as multimedia movie, my solution is, in fact, that the sense-of-self-in-the-act-of-knowing emerges within the movie, that it is a part of the movie, and thus creates, within the same imagetic space, both the "seer" and the "seen", both the "thinker" and the "thought". There is no need to posit a spectator for the movie-in-the-brain. The *idea* of spectator is constructed *within* the movie. This is a plausible answer to the vexing question of how the allegedly objective brain processes create the subjectivity of the conscious mind. There is no need for a homunculus in this solution of the conscious mind problem. The first person subjectivity is constructed from the ground up, based on the same kind of sensory mapping needed to construct the object representations made "conscious" by the sense of self.

A final comment pertains to the particular quality of the sensory map of self. Because the most fundamental sensory mappings pertain to body states and are imaged as feelings, the sense of self in the act of knowing emerges as a special kind of feeling, the feeling of what happens in an organism caught in the act of interacting with an object.

Concluding remarks It is certainly the case that we still cannot explain satisfactorily how the conscious mind emerges from the brain. But there is no principled hurdle blocking our attempt to answer the question to a substantial degree. To be sure we are not likely to elucidate the mystery fully but the resulting incompleteness does not appear to be of any greater magnitude than that with which science is faced in its attempts to understand matter in general. Specifically, the seemingly vexing hurdles of extracting a first-person "subjective" mind from a third-person "objective" body appears transposable with the help of the approach I advocate, that is, when the brain is seen as a complex device concerned with regulating and, of necessity, representing body states. I suspect Spinoza would have welcomed this solution and it is possible that Huygens might have been intrigued enough to write to him about it.

References

> Adolphs, R., Tranel, D., Damasio, A.R., 1994. Impaired Recognition of Emotion in Facial Expressions Following Bilateral Damage to the Human Amygdala, *Nature*, 372:669-672.

> Adolphs, R., Damasio, H., Tranel, D., Damasio, A.R., 1996. Cortical Systems for the Recognition of Emotion in Facial Expressions, *Journal of Neuroscience*, 16:7678-7687.

> Adolphs, R., Damasio, A.R., 1998. The Human Amygdala in Social Judgement, *Nature*, 393:470-474.

> Chalmers, D., 1996. *The Conscious Mind*. Oxford University Press, New York.

> Crick, F., 1994. *The Astonishing Hypothesis: The Scientific Search for the Soul*. Charles Scriber's Sons, New York.

> Damasio, A.R., 1989. Time-Locked Multiregional Retroactivation: A Systems Level Proposal for the Neural Substrates of Recall and Recognition, *Cognition*, 33:25-62.

> Damasio, A.R., Tranel, D., Damasio, H., 1989. Amnesia Caused by Herpes Simplex Encephalitis, Infarctions in Basal Forebrain, Alzheimer's Disease, and Anoxia. In: Boller, F. and Grafman, J. (eds.), *Handbook of Neuropsychology*, Vol. 3. Elsevier, Amsterdam: 149-166.

> Damasio, A.R., Tranel, D., Damasio, H., 1990. Face Agnosia and the Neural Substrates of Memory, *Annual Review of Neuroscience*, 13:89-109.

> Damasio, A.R., 1882. Aphasia, *New England Journal of Medicine*, 326:531-539.

> Damasio, A.R., Damasio, H., 1992. Brain and Language, *Scientific American*, 267:89-95.

> Damasio, A.R., 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. Putnam, New York.

> Damasio, A.R., Damasio, H., 1994. Cortical Systems for Retrieval of Concrete Knowledge: the Convergence Zone Framework. In: Koch, C. (ed.), *Large-Scale Neuronal Theories of the Brain*. MIT Press, Cambridge, MA: 61-74.

> Damasio, A.R., Grabowski, T., Bechara, A., Damasio, H., Ponto, L.L.B., Hichwa, R., 1998. Neural Correlates of the Experience of Emotions, *Society for Neuroscience*, 24:258.

> Damasio, A.R., 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Harcourt Brace, New York.

> Damasio, H., Grabowski, T., Tranel, D., Hichwa, R., Damasio, A.R., 1996. A Neural Basis for Lexical Retrieval, *Nature*, 380:499-505.

> Damasio, H., 2000. The Neural Basis of Language Disorders. In: Chapey, R. (ed.), *Language Intervention Strategies in Adult Aphasia, Fourth Edition*. Philadelphia, Williams & Wilkins/Lippincott.
> Dennett, D., 1991. *Consciousness Explained*. Little Brown, Boston.
Edelman, G, 1992. *Bright Air, Brilliant Fire : On the Matter of the Mind.* Basic Books, New York.
> Hubel, D., 1987. *Eye, Brain and Vision*. Scientific American Library. W.H. Freeman, New York.
> Kandel, E., Schwartz, J., Jessell, T., 1999. *Principles of Neural Science,* fourth edition. Appleton & Lange, Norwalk, Conn.
> LeDoux, J., 1992. Emotion and the Amygdala. In: Aggleton, J.P. (ed.), *The Amygdala: Neurobiological Aspects of Emotion, Mystery, and Mental Dysfunction*. Wiley-Liss, New York.
> Lopes da Silva, F.H., Witter, M.P., Boeijinga, P.H., Lohman, A.H.M., 1990. Anatomical Organization and Physiology of the Limbic System, *Physiological Reviews*, 70:453-511.
> Lopes da Silva, F.H., 1996. The Generation of Electric and Magnetic Signals of the Brain by Local Networks. In: Greger, R. , Windhorst, U. (eds.), *Comprehensive Human Physiology*, Vol. I. Springer Verlag, Berlin: 509-531.
> McGinn, C., 1991. *The Problem of Consciousness*. Oxford University Press, New York.
> Penrose, R., 1994. *The Shadows of the Mind.* Oxford University Press, New York.
> Searle, J., 1992. *The Rediscovery of the Mind.* MIT Press, Boston.
> Tononi, G., Edelman, G.M., 1998. Consciousness and Complexity, *Science*, 282:1846-1851.
> Tootel, R.B.H., Switkes, E., Silverman, M.S., Hamilton S.L., 1988. Functional Anatomy of Macaque Striate Cortex. II. Retinotopic Organization, *Journal of Neuroscience*, 8:1531-1568.
> Tranel, D., Damasio, A.R., Damasio, H., Brandt, J.P., 1994. Sensorimotor Skill Learning in Amnesia: Additional Evidence for the Neural Basis of Nondeclarative Memory, *Learning and Memory*, 1:165-179.
> Zeki, S., 1993: *A Vision of the Brain*. Blackwell Scientific Publications, Cambridge, MA.

Fernando H. Lopes da Silva

From Neuronal Networks to Consciousness

A Comment

Introduction The choice made by NWO of the theme Neurocognition for the Huygens Lecture is timely, in the context both of the international scene and of the Netherlands scientific environment. Internationally, the field of neurocognition exhibits a strong momentum. There is a growing interest for multiple aspects of research on the interaction between brain and cognitive functions. This enthusiasm is displayed not only by the scientific community but also by the public in general. The media has reflected this hype with a steady stream of information about neurocognition issues in the pages of reputed newspapers and weeklies, richly illustrated with colored images of the brain of subjects while executing cognitive tasks. A typical example is a very recent article of *Time Magazine* (Nov. 1st) with the heading "Telling right from wrong" suggesting that "scientists found our moral compass". Very appropriate to the discussion of today, this news was derived from an article of Damasio's group[1] where they report the study of two patients who suffered in early life severe injuries to their prefrontal lobes and thereafter were not able to develop a sense of right and wrong.

Remarkably renown scientists and philosophers alike, in the last decade, have dedicated thoughtful books to themes related to higher brain functions such as perception, attention, reasoning, memory, and consciousness. Among these Antonio Damasio occupies a very special position since he has been actively engaged, for a long time, in the daily observation of patients exhibiting the most florid panoply of neurological symptoms. He is a pungent observer of the workings of the brain of patients since the first hour of his M.D., at the same time sharp in his analyses and mentally stimulating in his interpretations. His career is a testimonial of the best tradition of neurology, based on careful clinical observation of the symptoms that the patients display allied with an insightful analysis of the processes underlying the clinical manifestations. Even before the new brain imaging techniques had emerged, his studies of language disorders, the theme of his Ph.D. thesis, were seminal. Making able use of his literary gifts, he gave us elegant accounts of his neurocognitive observations, thoughts and speculations in two most readable books[2,3].

Damasio was the appropriate person at the appropriate time, since he was able to incorporate the new advances in brain imaging with his already profound clinical insight. In this respect, the collaboration of the couple Antonio and Hanna Damasio was ideal, since Hanna brought into this field a solid knowledge of the techniques of brain imaging, as reflected in her notable *Atlas of Human Neuroanatomy from Computerized Images.*

At the opening, I mentioned that the choice of the theme for this Huygens lecture was not only timely in the international context, but it was also particularly pertinent for the situation in the Netherlands. Here the development of the field of the neurocognitive sciences has been relatively slow. Traditionally, the fields of the basic neurosciences and of cognitive sciences developed in parallel but at a respectable distance of each other. It is opportune to note that the last years a number of scientists from the two areas have made steady efforts to bring the two fields together in a joint endeavor to create new initiatives in the field of neurocognitive scientific research in the Netherlands. These efforts were fortunately well felt by NWO who created a thinking tank on this topic, to which I had the privilege of participating, in order to promote new initiatives in the research field of cognitive sciences. I hope that NWO will also be prepared to actively participate in the necessary investments both in minds and tools.

NWO asked me to give a commentary on the theme of Damasio's lecture The Fabric of the Mind: a Neurobiological Perspective. In formulating these comments, my perspective is thus that of a basic neuroscientist who wishes to express a few afterthoughts on "the fabric of the mind" and related enigmas.

Why brain areas should not be considered the "centers" of cognitive functions It is understandable that studies of the cognitive functions of human subjects with brain lesions can lead to the conclusion that the lesioned brain areas are responsible for the impaired cognitive functions. The classic view of the brain as consisting of defined brain areas with a strict functional correspondence has been recently reinforced by the generalized use of modern functional imaging techniques such as positron emission tomography (PET) and functional MRI (fMRI). The fact that certain brain areas appear to be preferentially active during the performance of certain cognitive functions, as apparent in brain scans, can contribute to forming a *static image* of the brain as consisting of series of areas identified as "centers" of cognitive functions, similar to what we see in the old pictures of the phrenology. In other words, the phrenology ghost looms in the background. Damasio also warns against this tendency. I could not agree more with his remark[2] about phrenology where he says that we should give credit to the founder of phrenology, the Viennese physician Franz Josef Gall "with the concept of brain specialization, an impressive idea given the scarce knowledge of his time, [but] we must blame him for the notion of brain 'centers' that he inspired"[2].

But if we should not think of the brain as a collection of "centers", what then? In my view, we should go further and consider the brain as consisting of *interlocked neuronal networks* which do not correspond simply to the classically defined anatomical areas. To substantiate this statement, I should advance two arguments: *first*, the fact that each cortical neuron is connected

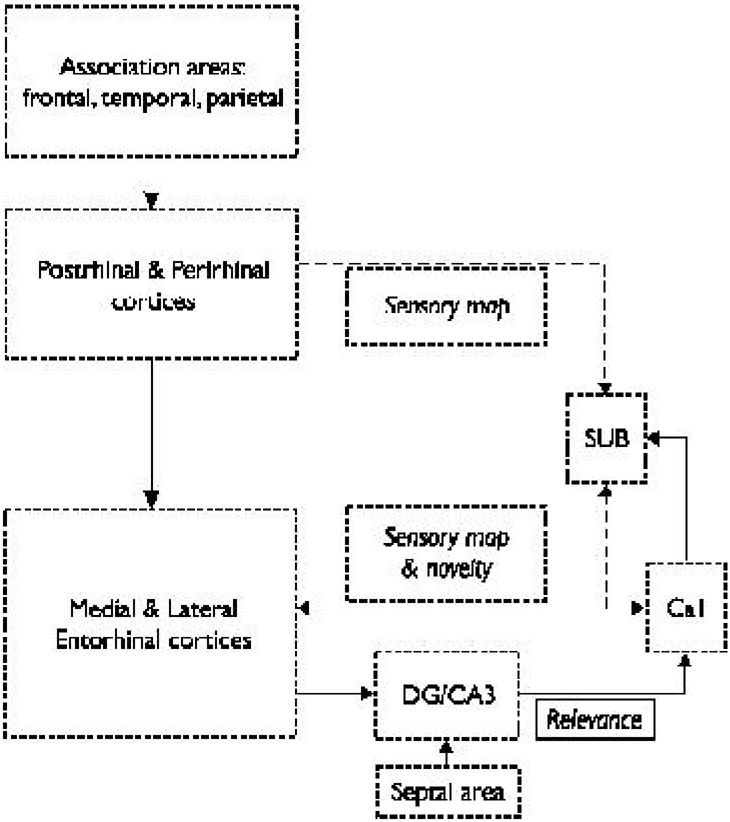


Figure 1: Example of a distributed assembly of neuronal networks of the temporal lobe including the Hippocampus (CA1 and CA3 areas and Dentate Gyrus-DG) and directly associated cortical areas (Lateral and Medial Entorhinal cortices, Perirhinal and Postrhinal cortices) that form a number of nested circuits where neuronal information is processed. The Subiculum (sub) occupies a pivotal position in this assembly of networks since it forms a convergence zone that receives associative sensory information of different modalities and distributes the processed information to several other cortical and sub-cortical brain systems. The associative sensory information reaches the Subiculum in three forms: (i) a raw version of the sensory map; (ii) a version that results from the comparison between new and stored information providing a "novelty flag"; (iii) a version weighted by the inputs arising from the neurochemical systems of the diencephalon and mesencephalon providing a "flag of emotional relevance". The hypothesis is that the Subiculum performs a dynamical comparison of these different versions and, accordingly, selects the information for storage and/or for action or just ignores it (Adapted from ref. 4).

to any other by at most three synapses implies that any interpretation of the brain as consisting of strictly circumscribed areas or centers is a gross oversimplification. *Second*, I would like to mention what the renown Dutch anatomist Hans Kuypers used to say, namely that one of the troubles of understanding brain functions has to do with how 19th century anatomists baptized brain areas with names that they coined on the basis of vague similarities between macroscopic brain structures and fruits, animals or common objects, like a pear (pyriform lobe), an almond (amygdala), a seahorse (hippocampus). Even today, a general tendency is to think that because a certain area has a given anatomical name, it should correspond to one function, i.e. *one anatomical name = one cognitive function*! However, such a strict correspondence does not hold in general. Of course this does not mean that there would not exist specialized brain areas that underlie specific functions, but this holds for rather simple functions such as those involved in the execution of motor behavior or in the primary processing of sensory information. For example let us consider the hippocampus. Scientists have been arduously pursuing many lines of research to unravel what *the function* of the hippocampus might be. Notwithstanding decades of investigations, this endeavor has not revealed a direct correspondence between the hippocampus and *one single* behavioral function. The main reason, in my opinion, is that the brain area which we call the hippocampus cannot be considered neither isolated from other brain areas with which it forms a number of interlocked neuronal networks, nor as a uniform structure that would constitute one functional unity. Recent research in our group[4] led us to the conclusion that the hippocampus forms with the parahippocampal region a number of discrete nested circuits where the flow of neuronal information re-enters the same brain areas after being submitted to different degrees of processing, such that these inter-locked networks evaluate whether a given sensory scene has the character of novelty or not, whether it is emotional relevant, or not, and then can pass the result of this evaluation to other brain systems where memories may be laid down (figure 1). In this context, these brain systems form a convergence zone in the sense used by Damasio[5]. In addition to the hippocampus, there are many other brain systems involved in memory functions, so that we cannot identify a site for memory in the brain, only that some brain systems are more specialized in the formation of particular aspects of memory than others.

These considerations lead naturally to the general *conclusion* that instead of identifying brain areas as the *centers* of cognitive functions, we should rather try to find out the organization of the *neuronal networks* that are responsible for such functions. Indeed a given brain area may be *necessary* for the execution of a certain cognitive function but not *sufficient* per se because it must interact with other structures in order to perform meaningful actions. In addition, the same *local* neuronal network may contribute to different macroscopic networks distributed across the brain, and thus it may participate in different cognitive functions.

Why the neurobiological substrates of cognitive functions should be considered to consist of dynamical neuronal networks Cognitive processes, such as the appreciation of a visual scene, the recognition of a melody or the identification of an odor, occur in a fraction of a second, i.e. they are highly dynamical processes. In this context the recent experimental observations of Roelfsema, Lamme and Spekreijse[6] here in Amsterdam are worthwhile mentioning. They showed, in the awake monkey, how the activity of a cortical neuron is associated with the fact that a given stimulus

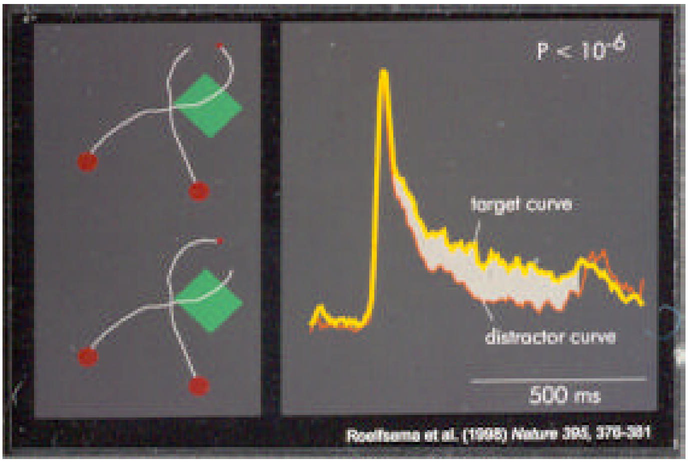


Figure 2: Experiment showing how object-based attention is associated with a response enhancement at the earliest (area V1) level of the visual cortical processing system of the monkey. The animal was trained, first to fixate on a spot displayed on a screen, and then to respond with a saccade from this spot to a circle at the other end of a curve; this is called the target curve; the other curve to which the animal should not pay attention was a distractor curve; the two curves intersected. Target and distractor curves were interchanged at given times. The cortical neuron the activity of which is displayed in the post-stimulus histogram, had a receptive field indicated by the square. Note that the responses (number of impulses per second is indicated along the vertical axis) to the presentation of the target and distractor curves differ significantly within a time window from about 60 to 500 msec. This indicates a neuronal event that is associated with the phenomenon of selective attention or conscious perception (kindly provided by Pieter R. Roelfsma, see also ref. 6).

receives selective attention, i.e. it is consciously perceived by the animal. The animals were trained to perform a curve-tracing task. The neuronal responses of a cell of the cortex were enhanced when the animal paid selective attention to a target curve relative to a distractor curve. This is a correlate of object-based attention. This occurs in a time window between 60 and 250 milliseconds after stimulus presentation (figure 2). This result shows how dynamic this process is, since from the time of stimulus presentation to the time window where conscious perception emerged, the neuronal information has traveled most certainly back and forth between several cortical and sub-cortical areas.

This example demonstrates that the neuronal networks responsible for conscious perception are essentially *dynamical* systems. However most classic neurocognitive studies are based on the analysis of patients with brain lesions or on brain scans with a poor time resolution. So, these studies necessarily miss the dynamical features which characterize *how* neurocognitive processes are executed. Many researchers nowadays hope that observations of such cognitive processes in patients or in normal subjects using PET or fMRI may furnish the missing information in this respect. As many others today, Damasio has made intelligent use of these new techniques. However there is a fundamental problem here.

The problem is that the time resolution of PET and even of fMRI is not adequate to grasp the *dynamics* of many neurocognitive processes. This problem is not generally realized. In this context, three main limitations of these brain scan techniques have to be taken into account: (i) first, the signals of brain activity which are used in PET and fMRI depend on changes in local metabolism and/or in vascular changes, and these occur with a time delay of at least 200 msec after the change of neuronal activity: this is a hard physiological limit; (ii) second, these signals have to be integrated over periods of seconds in order to be detectable using the current techniques so that the fine dynamics of the neuronal activity cannot be distinguished; (iii) third, the signals obtained from PET and fMRI are strongly biased in the direction of increases of activity, while the functioning of brain networks involves as much increases as decreases of neuronal activity. Does this imply that we will not be able to appreciate what happens in the brain during the short time window sufficient for a human being to consciously identify a visual scene or to feel the soft contact of a silk scarf? In a free analogy with cosmology, we may ask which experimental methods will ever be able to grasp the “big-bang” of conscious perception. This cannot be achieved by PET and fMRI techniques that can only give us signals of what happened in the brain after the essential moment of conscious perception; they are rather signals of the *aftermath*.

Searching for the coupling between transient changes in consciousness and in brain activity In order to catch the dynamics of cognitive processes as they unfold, it is therefore necessary to develop and explore new techniques. In this respect, the possibility of using whole-head magneto- and electroencephalography (MEG/EEG), in combination with fMRI, may provide the best practical solution available at present. MEG/EEG provide information with a high time resolution, in the millisecond range, although at the cost of a more modest spatial resolution than functional MR. The combination of both sets of techniques is starting to be developed, although there are still a good number of technical problems to be solved. We are currently exploring this approach in an investigation of transients in brain activity directly related to changes in the state of *core consciousness*, in the sense used by Damasio[3].

In this context our investigation concerns patients that suffer from a form of absence epilepsy triggered by a flickering light stimulus. These patients are exquisitely sensitive to intermittent light of the appropriate frequency. While viewing such a stimulus for a few seconds, the patients may suffer an impairment of consciousness, which is usually described as an absence, i.e. the patient usually stares blankly and his/her behavior seems to be in a frozen state without showing any emotion or communication with the outside world. Such an absence usually lasts only a few seconds. Afterwards the patient has no knowledge of what happened during this interval of time. At the same time an abnormal brain activity occurs which can be recorded by means of whole-head EEG/MEG. A practical advantage of investigating this type of patients is that the *absence of consciousness* can be externally triggered and, thus it can be studied under controlled conditions in a laboratory setting. The question that we ask is the following: which are the neuronal networks the working of which is affected by the abnormal oscillations such that consciousness is switched off. An answer to this question would give us some clues about the neuronal mechanisms that are essential for core consciousness. Our team of the Institute of Epilepsy in Heemstede, together with the group of the MEG-Center of the Royal Netherlands Academy of Arts and Sciences at The Vrije Universiteit in Amsterdam, is now carrying out such an investigation. The results are still preliminary. Combining MEG/EEG data with MRI, and using the appropriate computer methods of functional source analysis, we estimated that the occurrence of the abnormal oscillation, of low frequency (about 3 Hz), most likely involves mid-line neuronal networks that strongly interact with several brain cortical areas. This creates a state of increased synchronous oscillations throughout the brain, i.e. a hypersynchronous state. Under these conditions core consciousness is interrupted. This investigation is thus an example of studying the neuronal systems responsible for consciousness from the absence of the latter. The implication of these results is that under normal conditions, core consciousness is maintained by the dynamical interactions between several cortical and sub-cortical networks; these interactions are interrupted by the abnormal oscillation where thalamic nuclei play a prominent role under the control of upper brain stem systems. These mid-line networks are modulated by distinct neurochemical systems (acetylcholine, catecholamines, serotonin). These neurochemical modulators, jointly with humoral signals, cause changes in brain functions that are expressed as emotional states. The latter are important for the survival of the organism and correspond to the conditions of pleasure and pain, reward and punishment. Damasio puts forward that the expression of emotions precedes feeling, and that consciousness depends on feeling[7]. However at the present time, we are still short of achieving an insight of what feelings are really made of. I would hypothesize that feelings are generated at the interface between neuromodulator systems and networks responsible for core consciousness.

During the epileptic absence, the abnormal oscillation typically of low frequency (ca. 3 Hz) disrupts the normal flow of signals in these networks, i.e. it jams the working of these networks. We may say that as a consequence, core consciousness is temporarily switched off. It is interesting to note that in a normal state of impaired consciousness that all of us experience every day, that of deep sleep, the activity of the brain is dominated by low frequency oscillations, in the order of 0.1 to 4 Hz. A similar change in EEG activity occurs in deep anesthesia where consciousness is also impaired. Thus it appears that core consciousness is not compatible with low frequency hypersynchronous states, that involve the mid-line neuronal networks of the brain.

How does consciousness emerge from dynamical activity of neuronal networks? Consciousness is based on a well orchestrated organization of memories, since it implies a continuous flow of information relaying memories of events both external and internal, sensory or motor, emotional or visceral. This property is what Damasio calls “extended consciousness” in the sense that it “goes beyond the here and now of core consciousness, both backward and forward”[7].

Consciousness implies a neuronal system that keeps track of memories, including that of the own self, which are formed at different time scales. Such a system must keep track of a variety of neural representations on-line and achieve their integration into a unitary dynamical state. But how can this dynamic integration be accomplished?

One approach to answer this question is inspired by the experimental finding that the phenomenon of perception of a visual scene is reflected in the occurrence of synchronous firing of populations of neurons that tend to oscillate at relatively high frequencies, between 20 and 70 Hz. This forms the basis of the correlation hypothesis initially proposed by Milner and Von der Malsburg[8], later called "*binding hypothesis*" which assumes that temporal correlation between neurons of the sensory cortex would form the basis of the phenomenon of unitary perception[9]. Extrapolating to the domain of core consciousness, I may hypothesize that the dynamical binding of neuronal activities occurring at relatively high frequencies (beta/gamma range) across several neuronal networks of the core of the brain may be a condition for the emergence of consciousness. According to this hypothesis the slow frequency of the hypersynchronous state characteristic of the epileptic absence discussed before, would impair the normal flow of the high frequency activity necessary for the "tracking system" underlying consciousness to work normally. The question now is to obtain direct experimental proof for this hypothesis.

As Damasio states, "consciousness results in enhanced wakefulness and focused attention"[7]. Along this line of thought, an experimental finding appears to be relevant. Namely that some cortical areas tend to show beta/gamma (30-40 Hz) EEG/MEG oscillations in close association with focused attention when a subject is asked to make a voluntary finger movement, while oscillations at lower frequencies (about 10 Hz) are depressed. This necessarily corresponds to increased correlated activity of cortical neurons in the beta/gamma frequency band[10]. In conclusion there is evidence that changes in the oscillatory dynamics of cortical neuronal networks are associated with processes of focused attention that are part of core consciousness.

Some researchers, however, doubt that experimental approaches may yield useful results due to the complexity of the neuronal systems and to technical difficulties. As an alternative they put their hopes in "connectionist theories" in order to unravel the mechanisms of "the fabric of the mind" and thus of consciousness. Their viewpoint is that machines can be constructed which may have such a degree of complexity, that their intelligent performance may not be distinguishable from that of a human being and therefore they may be said to have consciousness. Consciousness would be an emergent property of the activity of dynamical connections in such complex networks. Is this likely to occur?

This line of thought has already a long tradition. The question of whether computers could help to understand the more complex functions of the brain was first formulated by Alan Turing[11] who wrote, back in 1950 an influential essay entitled “I Propose to Consider the Question *Can Machines Think?*”, i.e. can machines be built that have the property of consciousness as the brain has? Since then this question has been a source of controversies.

For some authors, like Aleksander[12], the answer appears to be positive. Aleksander admits that the currently most powerful computer, called Magnus and developed by his group at the Imperial College of Sciences in London, does not in itself, as yet, possess the features of consciousness, yet he assumes that Magnus might develop consciousness in the future given the advancements of computer technology. Of course this is more an argument of faith than one with scientific validity. Nevertheless connectionist modeling may be useful to simulate some kinds of brain operations in such a way that these models may help to formulate testable hypotheses about the neuronal mechanisms underlying cognitive functions. Anyway the fact that a given brain function may be simulated in a computer does not mean that the computer *possesses* such a property, similarly to the situation that a simulation of a weather system in a computer does not mean that the computer is capable of creating the weather, although this may be useful to predict whether it will rain or not. Be as it may, the pretension that connectionist modeling might account for consciousness is still far-fetched. The main shortcoming of this kind of modeling may be what Damasio convincingly stressed by writing that human consciousness requires the existence of feelings, and “what feelings feel like cannot be duplicated in silicon chips”[7]. I agree that at this moment it is difficult to imagine how this would happen but Turing’s question remains unanswered.

In conclusion we may state that recent contributions, both at the theoretical and experimental levels, are unraveling which neuronal networks in the brain are necessary for the execution of some cognitive functions, including consciousness, and also how these functions may emerge from the underlying neuronal networks. Nevertheless new technical advances are necessary in order to get a much more concrete idea of the basic mechanisms that occur in a neuronal network, making it possible in a few milliseconds for consciousness to emerge. Francis Crick[13] wrote that “consciousness is now largely a scientific problem. It is not impossible that, with a little luck, we may glimpse the outline of the solution before the end of the century”. We are certainly getting closer to developing a deeper insight into the problem, but the end of the century is also pretty close.

References

[1] Anderson, S.W., Bechara, A., Damasio, H., Tranel, D. and A.R. Damasio, 1999. *Nature Neuroscience* 11: 1032-1037.

[2] Damasio, A.R., 1994. *Descartes' Error*. Grosset/Putman, New York.

[3] Damasio, A.R., 1999. *The Feeling of What Happens*. Harcourt Brace, New York.

[4] Naber, P.A., Witter, M.P. and F.H. Lopes da Silva, *Annals New York Academy of Sciences*. In Press.

[5] Damasio, A.R. and H. Damasio, 1994. In: C. Koch and J.L. Davis (eds.) *Large-Scale Neuronal Theories of the Brain*. Cambridge, Mass, MIT Press: 61-74.

[6] Roelfsema, P. Lamme, V. and H. Spekreijse, 1998. *Nature*, 395: 376-381.

[7] Milner, P., 1974. *Psychol. Rev.*, 81: 521-535.

[8] Von der Malsburg, C., 1995. *Curr. Opin. Neurobiol.*, 5: 520-526.

[9] Singer, W., 1993. *Annu. Rev. Physiol.*, 55: 349-374. Gray, C.M., 1999. *Neuron*, 24: 31-47, but see criticisms in Shadlen, M.N. and J.A. Movshon, 1999. *Neuron*, 24: 67-77.

[10] Pfurtscheller, G. and F.H. Lopes da Silva, *Clinical Neurophysiology*, 110: 1842-1857.

[11] Turing, A.M, 1950. *Mind*, vol. LIX, no. 236.

[12] Aleksander, I., 1996. *Impossible Minds: My Neurons, My Consciousness*. Imperial College Press, UK.

[13] Crick, F., 1996. *Nature*, 379: 485-486.