Cognitive Neuroscience and the Hard Problems

In the philosophy of cognitive neuroscience, the problem concerning the obvious differences between mental states and brain states stands out as more stubborn than any other. The situation seems analogous to a situation within philosophy of physics. In physics, we see how hard it is to merge quantum theory and the general theory of relativity into a single unified theory. This is partly due to the fundamentally different approach each of these theories takes to the world. Quantum theory views the basic constituents of the material world as made of discrete corpuscles, whereas the theories of relativity and classical electromagnetism regard these constituents as continuous fields. In addition, we find within quantum mechanics itself the particle-wave duality, where quantum objects sometimes behave as if they are waves, sometimes as if they are particles, all depending on the experimental arrangement. Finding a solution to these unavoidable particle-wave dualities belongs to what we might call the hard problem of modern physics.

We find a similar dual challenge in cognitive neuroscience. Here the nature of our experiences seems fundamentally different from the nature of the underlying brain processes. This hard problem arises, apparently, because of the conceptual dichotomy between the quantitative approach followed by neuroscience and the qualitative approach informing introspective and much experimental psychology. In addition to the qualitative-quantitative contrast, we shall identify and discuss three other enigmas as hard problems with respect to cognitive neuroscience. These are the mentality problem, the consciousness problem, and the self-consciousness problem.

Almost every philosopher and cognitive neuroscientist believes that the brain somehow creates the mind, but nobody really knows how this can happen. On the one hand, neuroscience focuses on the interacting nerve cells of the brain as the realizers of the mind. On the other hand, ordinary psychology and cognitive science attribute qualitative sensations, experiences, reasoning, and intentions to the conscious mind. The challenge arises because, in general, philosophers and scientists agree that not only has nobody yet understood how sensory experiences, or consciousness for that matter, are explainable in terms of a physical-chemical description of the activity of neurons. But some even believe that cognitive neuroscience will never be able to come up with a satisfactory solution to that problem.

Recently, Peter Carruthers has argued that there are no conscious thoughts. His claim is not as radical as it sounds. He only argues that *episodic* thoughts such as judgments and decisions take place outside consciousness; a conclusion he reaches by considering consciousness as defined in terms of the global broadcasting model or in terms of the non-interpretative higher-order awareness. He does not rule out that we may be conscious of our sensory experiences or non-episodic thinking (Carruthers 2017).

The confusion behind much of the ongoing discussion is partly a result of the fact that common terms like "mind", "consciousness" or "conscious mind" are all ill-defined and often used on an intuitive basis. Some divide "consciousness" into "creature consciousness" and "mental state consciousness"; others talk about "access consciousness" in contrast to the unconscious part of the mind that is inaccessible to mental scrutiny, and others again focus on the phenomenal consciousness as the subjective side of cognition that falls outside scientific investigations. The reason why the central mental terms are so ambiguous is that we have only direct access to our cognitive processes. We are aware of our own thinking (internal voice), our own knowledge, our own memory, or our own mental images. Hence, we cannot imagine how it is to think, know, remember or imagine without being aware that one is actually engaged in one of these activities.

However, if we are ready to say that animals other than humans are cognitively adapted to be aware of their surroundings, it seems reasonable to claim that these animals can be conscious about certain things in their environment. Of course, such an ascription of consciousness to many species of vertebrates requires that we can conceptually distinguish between three different categories: organisms capable of reacting to sensory stimuli,

organisms capable of being aware of external circumstances, and organisms not only capable of being aware of external circumstances but also introspectively aware of some internal mental states. Elsewhere, I have argued that *mentality* is the overarching concept and that *primary* and *secondary consciousness* are subcategories for characterizing the difference between the awareness of external objects and the awareness of having certain mental states oneself (Faye, 2019). Roughly, this classification divides cognitive beings into sentient creatures, thinking animals, and self-reflective animals, although evolution reminds us that there probably exists a continuous spectrum of species in between.

As I see it, mentality consists primarily in the ability to have sensations and to learn from these sensations. If this is correct, sensations by which an organism senses something might be "subjectively" different from the physical stimulation of the sense organs. Consequently, in order to explain the phenomenal aspect of sensation, I shall propose a solution to this problem by making a distinction between *intrinsic* and *extrinsic* properties. The argument is that we have to describe all systems, regardless of their complexity, not only in terms of their intrinsic properties but also in terms of their extrinsic properties, which these systems have whenever they interact with their surroundings. In relation to cognitive neuroscience, the brain has intrinsic properties that we attribute to neurons and the other functional subsystems in the brain, whereas the extrinsic properties of the subsystems are those properties the brain acquires through a causal interaction with its environment. These extrinsic properties, I argue, constitute our mental experiences and our thoughts about them.

The aim of cognitive neuroscience

Neuroscience has managed to analyze the function of some of our cognitive abilities and emulate these functions based on algorithmic simulations of this neuronal activity. And many more discoveries will follow. But neuroscience has little to offer about the nature of the mental other than what it does for an organism. There seems to be a subjective side of the mind for which it cannot account, due to a particular feel claimed to exist in connection with having specific experiences. In particular, Thomas Nagel has characterized this as the sensation of what it is like to experience something (Nagel 1974). This is one of the great tests for merging neuroscience and cognitive science. How can cognitive neuroscience explain what it is like to experience red, taste something sweet, or feel pain? This is *the qualia problem*, which Chalmers originally doubted the hard problem (Chalmers 1996).

Regardless of the qualia problem, cognitive neuroscientists seem to face an even more fundamental problem: how is it possible to understand the content of our experiences and emotions, which we usually characterize in terms of qualitative properties, exhaustively in terms of quantitative properties associated with brain processes. This general problem I call *the quality-quantity problem*. We shall characterize a qualitative property as a quality that we attribute to an entity based on our sensory experience of that entity. In contrast, a quantitative property is one that has a distinct and measurable value and is ascribable to an entity based on theoretical considerations. In other words, the quality-quantity problem concerns whether or not we can explain those qualitative properties associated with what we experience, feel, imagine, or think by an appeal to the quantitative properties by which neuroscience describes the brain.

Indeed, few philosophers, like Patricia and Paul Churchland, reject the existence of qualitative properties altogether by arguing that such an assumption is part of an old-fashioned psychological theory that a matured neuroscience can ignore without any significant cost. Still, not many philosophers believe that an eliminative approach is cost-free (Churchland, 1981).

Among philosophers, psychologists, and neuroscientists other reactions to the quality-quantity problem come to the fore. One response, first made by Thomas Nagel, is to argue that, at least at present, science is unable to come up with a satisfactory explanation (Nagel 1974). This is the agnostic approach. The reason for such an agnostic view is that the current cognitive neuroscience lacks the conceptual resources to cope with qualitative descriptions in terms of quantitative ones, and our general understanding of the mind is essentially qualitative in

nature. A satisfactory account of the mind is possible, only if we can see how the conceptual description expressed by the explanandum translates into the conceptual description expressed by the explanans. Such a translation is not within our reach with respect to the understanding of the mind in terms of current neuroscience, but nobody knows what might happen in the future.

However, one could take a more skeptical attitude. One might argue, as Colin McGinn does, that such a translation is impossible, owing to the fact that we do not possess the cognitive powers by which we are able to recognize the correct theory, which would provide us such a translation (McGinn 1989). Alternatively, one could argue that mental properties do not have the same nature as physical properties. The first view is a purely epistemic one, whereas the second view relies on certain ontological considerations.

Take the epistemic view first. It assumes that there is a correct natural theory that can explain how qualitative properties rise from quantitative properties in a natural world, but we are cognitively unable to discover the content of such a theory. The argument for this view is straightforward: quantitative descriptions and qualitative descriptions are so different in meaning but so embroiled in our way of comprehending the world that it excludes us from finding the correct theory by which we can cross the explanatory gab. Accepting this argument, one could rightly submit that the conclusion does not follow, namely that there is one correct theory that explains the dichotomy between the qualitative and quantitative conception of the world, although the true theory is unrecognizable by us. Who can say that there exists such a unifying, but intellectually unintelligible theory, which brings apparently incoherent concepts together?

Another possible answer is that both descriptions provide us with an insight, which the other description cannot deliver, but that they together exhaust all there is to know about the mind-body problem. Rather we need a number of different theories to comprehend mentality. No qualitative expression can replace a quantitative expression, or vice versa. The difference in meaning reflects the difference in the situation of seeing things from a first person perspective and from a third person perspective. No explanatorily powerful theory can eliminate one of these perspectives in favor of the other. Although this answer may be true, many people will probably consider it incomplete. What remains to be shown is how those two kinds of descriptions can be given an ontological underpinning.

One way of providing an ontological foundation of the dichotomy between qualitative and quantitative descriptions is to declare that the mental and the physical consist of two distinct sets of properties attributed to the same underlying neutral substance. However, as a metaphysical position, property dualism faces serious challenges (Chalmers 1996). First is the question about where to stop attributing mental properties to substances. Panpsychism does not have to struggle with this challenge, because it ascribes mentality all the way down. However, as soon as property dualism attempts to make a distinction between entities that instantiate mentality and entities that don't, such a distinction seems to rely on physical consideration. Mentality becomes a question of how physical complex a system is. We are no longer talking about only mental attributes of a neutral substance, but about mental attributes in correlation with physical properties characteristic for a particular kind of physical substance. Property dualism also has to account for why these correlations seem robust and how they originate.

Assuming that no causal connection exists between mental and physical properties, it is impossible to see how robust correlations of this sort can continue to exist. On the other hand, if we assume that it is possible to account for these correlations in terms of some causal connections, the proponents of property dualism have to explain in detail how such connections are possible without being guilty of the same objections posed against Descartes' substance dualism. Property-dualism is a viable ontological possibility only if one can be certain that any entity that shares both physical and mental properties shares these properties by law and not by accident. Moreover, the position has to make sure that it does not collapse into epiphenomenalism by establishing that mental properties, in the form of qualitative experiences, takes an active part in the causal ongoing between the brain and the mind. These two requirements are strictly connected, but as far as I know not yet met in any acceptable way.

Another approach to resolving the question of ontological underpinning is to see qualitative descriptions as referring to mental experiences as epiphenomenal. This view seems to underlie much research within cognitive neuroscience. The advantage of such a strategy is that every part of the cognitive process that is not explainable in purely neuro-chemical terms is excluded from a causal description of the world. All significant processes are physical interactions among the neurons. Therefore, the non-causal residue of a cognitive process, if it exists as a subjective appendix, may be said to be an epiphenomenon and is, for all intents and purposes, explanatorily irrelevant.

However, this strategy is doubtful for a number of reasons, too. First, it is impossible to understand how we can know that qualitative sensations exist and react physically on this knowledge if qualitative sensations have no causal influence on the brain processes. Therefore, if we do know that qualitative sensations exist, they cannot be epiphenomenal. Second, if we make a reasonable assumption that qualitative sensations have been subject to adaptation and natural selection, just as the evolution of the brain has, it is inexplicable why qualitative sensations should be epiphenomenal. As an epiphenomenon, a qualitative sensation would be of no use to the organism, and the capacity of instantiating them would therefore not undergo continuous selection and adaptation. However, looking at the biological evolution it seems to vindicate that the more complex the nervous system has become, the more complex and sophisticated are the mental properties attributed to the involved organism.

Functionalism and neuroscience

The methodological foundation of cognitive neuroscience is functionalism. The purpose of this approach is to avoid wondering about the nature of consciousness and instead focusing on the function of various cognitive elements in connection with the neuronal activity of the brain. Neuroscience fulfills its scientific goals if it can explain how the brain processes external and internal information for the purpose of the organism's proper reaction to the content of this information. Eventually, there would be no scientific reason to understand mental states in other than functional terms, because this makes possible the explanation of both their evolutionary history and their causal role in establishing the capacities of the mind. Nonetheless, functionalism also comes in different versions that seem to reflect the above-mentioned quantitative-qualitative dichotomy.

One version is machine functionalism. It attempts to understand the activity of the neurons as governed by an innate algorithm of computations. The assumption is therefore that an advanced computer is able to simulate the activity of mental processes. Whatever the status of neuroscience is, the aim of this version is to give a quantitative description of the working brain. Adherents of this version sometimes claim that if it is possible to make a simulation of the various functions of the brain, such a simulation will bring mental processes to life and make computers conscious. A reasonable objection is that considering other cases of computer simulations, like e.g. of hurricanes, suggests that simulations of a particular system does no bring the system to life.

The other version is causal functionalism. It is much closer to the thinking behind cognitive science because it focuses on psychological states as qualitatively individuated. This version identifies every mental state with a certain state having a specific causal role in relation to the behavior of an organism. The main philosophical challenge that confronts cognitive neuroscience is therefore whether these two functional approaches are unifiable into a single theory. Is it possible to resolve the quality-quantity dichotomy?

Apparently, functionalism in itself does not give us a satisfactory methodological response to the dichotomy in question. The qualia problem deflects a functional analysis – or so it is argued. But one may refuse to accept *qualia* as more than a subjective sign of being in a particular experiential state in contrast to experientially observing that state, and therefore to reject that the qualia problem challenges cognitive neuroscience. Thus, while cognitive neuroscience may be relieved from the qualia problem, the quality-quantity problem is still there for it to solve. And apart from the quality-quantity problem, functionalism has difficulties explaining two or three other issues concerning mentality and consciousness.

Here are the hard problems – liberated from the issues of the subjectivity of qualia – which cognitive neuroscience cannot successfully solve only based on a functionalist methodology.

The mentality problem: what makes a physical state P a mental state rather than a non-mental state?

The quality-quantity problem: what makes a physical state P a mental state of sensing red rather than one of a state of sensing blue or sensing yellow?

The awareness problem: what makes an organism conscious of what it experiences rather than non-conscious?

The self-awareness problem: what makes an organism conscious that it has experiences?

All four questions pose a hard problem with respect to cognitive neuroscience because they are concerned with the nature of mental states and not their function. As long as cognitive neuroscience is unable to explain the nature of mental states, there will always be disagreement about what kind of insights cognitive neuroscience is able to provide. It is mandatory of cognitive neuroscience to open for a broader ontology that includes an account of mentality and consciousness, in case it does not want to leave parts of the mind-body problem in explanatory darkness.

Indeed, one may argue that the nature of mental states is a philosophical question, which the sciences do not need to address. The nature of mental states goes far beyond what we can empirically establish; thus, it is a metaphysical question more than a scientific none. To a certain extent, this is true, but every matured science embodies an ontology, which scientists within a field accept as part of their theoretical background. Unless cognitive neuroscience presents nothing but an instrumental perspective on neuronal processes, we need an account of the nature of mental states that accords with what we know about evolutionary biology and psychology in particular.

The mentality problem

A first tentative step is to ascribe mentality to a physical system in terms of its capacity of using external information to present what this information is informing the organism about. A system possesses mentality only if it is adaptively disposed to present something outside of itself for the benefit of its own behavior and it is adaptively disposed to learn from how the actual presentations turn out. This ability seems to be closely associated with the evolutions of neurons. Mentality seems to be a feature of neuronally equipped animals. It appears due to an internal construction of their environment based on external stimulations and enables them to behave according to such a construction. However, this does not give us the full answer. We want to know both what presentation means and whether presentation in nature confines to only biological organisms.

In a social and cultural context, a representation is the result of an act of will where we intentionally assigned some physical elements with a representational function. A representation is in such cases an *artificial construction*. A representation designed by humans is a system that carries information about another system in the sense that the representational elements of the system stand for something selected by a human being. This typically holds for maps, pictures, mathematical models and natural language sentences; whereas a sketch looking like president Trump does not represent him if the similarity is purely accidental. We have a representation only if intentions have designed a system such that it contains some information about features of another system.

If there is no intention involved, we may talk about a *presentation* rather than a *representation* as a *natural construction* of the environment designed by natural selection. Many natural systems contain information about other systems, although some of them are not adapted for collecting information for beneficiary purposes. The

requirement for being a natural presentation seems not only to be that the system holds information about features of another system. The information must help the organism to reproduce and survive.

Take a tree standing at the shore of the ocean. The shape of the tree may count as evidence for the fact that the wind normally blows from the ocean, the position of the leaves informs us about the direction of the incoming sun light, and the color of the leaves carries information about the season. This kind of information does not provide us with a presentation. Moreover, it is not only living beings that might have information about their environment without this information function as a presentation for the organism. For us, perhaps, this information represents evidence. In the same way as the position of a compass needle is evidence for the magnetic North Pole and the Uranus' orbit is evidence for the existence of Neptune. However, because there is no implicit reference to human intentions in these statements, I prefer to say that the compass needle *presents* evidence *to us* for the direction of the magnetic North Pole, and the Uranus' orbit *presents* evidence *to us* about the existence of Neptune. In order for a natural system to function as a presentation there must be some organism that can hold the content of the relevant information.

Besides this minor point, a more important one stands out. Since most natural systems can carry or react to information about their environment, mentality is not identical to information stored in a system. We must therefore see what may distinguish sentient beings from other organisms, like plants, not to mention physical systems in general. The distinction should tell us about the difference between information presenting something to the system on the one hand and effects merely being caused by something on the other.

A presentation is a biological construction. Psychology and cognitive neuroscience have shown repeatedly how a visual presentation does not mirror the actual size, shape, or color of the object presented to the organism. Information processing mechanisms transform information that comes in through the sense organs and sediments into quite different forms of information. These later forms of information constitute a presentation. The presentation consists of form and content and both are a result of a biological construction. The form is the kind of sensation by which a presentation presents the object of presentation, and the content is the presented object as its information carried by these sensations. We may think of a sensory experience as consisting of certain kinds of sensations that to the system function as a presentation. Visual sensations, for example, function as a presentation because an organism has adapted to treat the information they carry to be about objects external to its body. Hence, presentation is processed information treated by an organism's biological system as providing evidence for the object of presentation.

The distinctive feature of sentient organisms is that they contain neurons. Neurons are nerve cells adapted to receive, process, and store information in order to help the organism behave according to the content of this information. This construction presents the environment to the organism. Hence, the answer to the question of why some presentations are mental rather than non-mental is that mental presentations is an *extrinsic* property we find in organisms in virtue of having a neural network that helps them to survive and prosper. Plants also act on external stimulation; but, apparently, they need no neurons to survive. They manage to stay alive without, because they do not change location, and thereby environment, by moving around. There is nothing for them to learn. In contrast, animals can react back by changing location, and for such a capacity to be beneficiary more complex organisms must be able to store and process information that can assist them avoiding dangerous situations and recognizing safe conditions, distinguishing food from non-food, or discerning rivals from mating partners. In other word, an organism that responds to its environment by changing location will benefit enormously from mechanisms that allow it to distinguish between different environmental stimulations and to remember the different stimuli, as well as let it react flexibly to the various inputs.

Summing up the claims so far, we can say that sensory experiences are presentations, which the neurons of the brain produces whenever they receive appropriate external information, in order to help the organism to behave successfully in relation to its habitat. However, no presentation is reducible to explanations in terms of internal properties of neurons, because not only can the content of a presentation not be individuated and classified

unless it is understood in relation to what presentation is about, but as we shall see, although physical in nature, the presentation consists of some extrinsic properties of the brain.

The quality-quantity problem

Perhaps the greatest enigma of all. Why do we experience color, sounds, smells, tastes and pain when *sensations* mirror neither the world of physics nor the physical-chemical properties of neurons? A physical description of our perception of a physical object does not attribute qualitative properties to the object we experience. Nor does a detailed description of the intrinsic properties attributed to neurons find any room for the qualitative properties attributed to sensation. We all know perfectly well that the neuro-surgeon will never find anything like red or blue sensations while looking into a brain. This has given rise to the qualia-problem in the philosophy of the mind and numerous philosophers and scientists have suggested that this question is out of scientific reach unless they, like Daniel Dennett (1997), have "quined" the qualia.

Here are some ideas that may move the discussion forward. Further details can be found elsewhere (Faye 2019). All systems are surrounded by other systems with which they are able to causally interact. These other systems constitute a given system's environment. Any given system is characterized by some internal states that are specified by a set of *intrinsic* properties. This holds regardless of whether we are talking about atoms, neurons, brains or stars. The intrinsic properties provide the appropriate system with certain dispositions, i.e. the abilities to interact with and react to the system's environment. When the environment activates these dispositions, it happens because the environment causes some changes in the system. These changes are instantiations of extrinsic properties. Through interaction with its environment, the system acquires a set of *extrinsic* properties that exist only as the result of the causal changes the interaction creates in the system. Thus, the manifestations of a system's dispositions, capacities or abilities always takes place in virtue of the system's environment that reshapes its internal states such that it exhibits some extrinsic properties that are either irreducible or unexplainable in terms of the intrinsic properties.

If one accepts this analysis, one realizes that the nature of a mental presentation is nothing but a set of *extrinsic* properties of the involved brain that it exhibits whenever the sense organs are stimulated by external influence. Although brain states are characterized in terms of the *intrinsic* features of the neurons, mental states such as sensations, memories, and images are *extrinsic* features associated with a certain set of neurons as they become activated by other neurons and by external stimulations. Hence, sensations of colors, odors, tastes or pains are extrinsic properties that appear as the final result of a long chain of information processing mechanisms in the brain, beginning with external activation of the receptors and ending up producing a presentation that involve many neurons in the brain. Thus sensing qualitative properties is a result of biological adaptation. Qualitative properties are those extrinsic properties that an organism is adapted to produce under the influence of the physical stimulations from its environment. Sensing qualities qua qualities has survival value.

Some might think of the extrinsic properties as *emergent* properties in disguise, but this I think is a wrong interpretation. We are not talking about a new, unexplainable level of properties that appears because the system is more complex than its elements. Emergent properties are intrinsic properties of a system, which the system has because it is this type of system, not because it interacts with a particular environment. In contrast, extrinsic properties are the outcome of a systems causal interaction with its environment. The existence of these properties depends just as much on the nature of the environment as on the nature of the system itself.

The awareness problem

Some philosophers argue that mental representations is what characterizes consciousness. For instance, Fred Dretske holds that if an organism represents something, it is conscious of what it represents (Dretske 1995). If we embrace this view, we should attribute consciousness to every sentient beings. However, the function of a

nervous system seems to be its capacity of presenting a habitat to the organism in such a manner that it can learn to behave in certain ways from the acquired sensations. The evolutionary benefit of a sensory experience is, I suggest, that it contains condensed information with respect to what the environmental stimulations contain. In relation to stimuli information, a nervous system reduces the amount of information needed for the organism to absorb in order to conduct successful behavior. Overall, a sensory experience should be understood as the outcome of a causal transformation of the stimuli-information.

An organism like *Caenorhabditis elegans* contains a little more than 300 hundred neurons, but it is able to learn and memorize what it has learned (Ardiel & Rankin 2010). It is not a hard-wired automaton. It reacts differently to different physical stimulations, and it learns which to avoid and which to approach. Assuming that the *elegans*' small amount of neurons helps it to process stimuli-information into various sensations of the world, and thereby makes it capable of constructing a sensuous presentation of its environment, does this acknowledgement force us to think of it as conscious?

A reasonable answer is no. Reasonable or not, an answer to how we *shall* understand consciousness depends very much on the agreed definitions and less so on reality. Another such understanding is to introduce higher order cognition. When we talk about consciousness, we often mean that consciousness involves an inner attention; our consciousness focuses on our mental states. For example, we can perceive things and phenomena around us, or feel pain, without paying attention to perception or pain itself. This fact has made David Armstrong (1970) claim that consciousness should be understood as perception of our own mental states. Thus, we are talking about the fact that consciousness is a higher-order perception of a first-order perception (HOP). An organism is conscious in case parts of its mental setup are able to monitor other parts of its internal life. The problem with that suggestion is that it does not explain the phenomenal content of our first-order perceptions. Moreover, and more importantly, if we are not conscious about the content of first-order perceptions before we perceive them, what then makes us conscious about the second-order perception? And why are we not automatically conscious of objects of our first-order perceptions, if we automatically become conscious of our mental states that are objects of a second-order perception? What makes the difference?

Instead of Armstrong's perception model of consciousness, David Rosenthal proposes a different model for higher order cognition, which seems to avoid some of these problems. First, he rightly points out that we cannot judge what it is like to be in a mental state of which we are not conscious: "There is, of course, nothing it's like to have a pain or a sensation of red unless the sensation in question is conscious. And some have argued from this to the conclusion that sensory quality simply cannot exist unless there's something it's like to have it. But what it's like for one to have pain, in the relevant sense of the idiom, is simply what it is like for one to be conscious of having that pain. So there won't be anything it's like to have pain unless the pain is conscious" (Rosenthal 2002, 411-12). In other words, according to Rosenthal, consciousness is the primary concept. We can only understand how pain seems to us in case we understand what consciousness is, because nothing seems to us if we are not aware of it. So we cannot use the concept of "what something is like for us" to explain what consciousness is. However, if we can explain what consciousness is, then we can explain why something seems to us in a particular way.

Second, Rosenthal makes the classical distinction between creature consciousness and mental state consciousness. Creature consciousness is present when an organism is awake, active, and exhibits behavior, whereas mental state consciousness is present when an organism is conscious of its own mental states. Then he makes another distinction between *transitive* consciousness and *mental state* consciousness: "… we understand transitive consciousness – being conscious of things – independently of understanding what it is for mental states to be a conscious state. We are transitively conscious of something by virtue of being either in an intentional or a sensory state whose content is a distinct property from that of a state's being conscious" (Rosenthal 2002, 407). However, our mental states are conscious because we actually have a thought about them (HOT).

Higher-order thoughts must not be confused with introspection: "A state is introspectively conscious when the addition is HOT. Ordinary, non-introspective state consciousness, by contrast, occurs when the HOT is not itself conscious" (Rosenthal 2002, 410). So in many cases HOT is not a conscious state, but in cases where HOT

is a conscious state, we have to deal with introspection.

Thus, Rosenthal sees consciousness as a higher-order thought (HOT). We are conscious of a mental state, like a sensory state, because we have a thought that we are just in this state. Being objects of thoughts, mental states become conscious. Some mental states in the organism monitor other mental states and make these states conscious. We avoid an infinite regress because there are many mental states that are not conscious, including our higher order thoughts that are not subject to a thought of an even higher order.

Both HOP and HOP fit well with functionalism, because first-order perceptions might be understood by their causal role, and the internal monitoring of first-order perceptions could therefore be described functionally. These two models also fit with access consciousness, because the monitoring allows the content of the mental state to be available to other cognitive activities. Moreover, they are acceptable with respect to what we know about biological evolution, since consciousness can be seen as a cognitive function that helps an organism to survive by coordinating all sensory impressions from the environment. But none of these models seems to fit particularly well as an explanation of phenomenal consciousness. Why do we experience red things as red in the first place?

Nevertheless, HOT suffers from serious shortcomings as a scientific model of consciousness. First, as others have noticed, too, sensory qualities are much more abundant than of what we have concepts. Second, HOT seems to deny consciousness to young children and non-human animals. Third, perhaps the most serious fault of all – one it also shares with HOP – it attempts to explain how one mental state becomes conscious in terms of another mental state. No higher order theory is actually able to bridge the gap between cognitive science and neuroscience. Overcoming the hard problem concerning awareness should not seek to explain the mental qua the mental but qua the physical. What perceives a first-order perception or has a higher-order thought? Uttering a sentence like "I feel pain," does "I" refer to *my* second-order perception, *my* higher-order thoughts, or to me as an organism that has this second-order perception or a higher-order thought? A long tradition points to a personal self. But is this self-constituted by perceptions of perceptions or thoughts of thoughts?

Self-awareness problem

We still have to explain how mental states become conscious before we can answer the self-aware problem. If the conscious mind cannot be explained by an appeal to other mental states like second-order perceptions or higher-order thoughts, we must reject the monitor-analogy according to which one mental entity surveys another mental entity. A reasonable answer to what consciousness is must be fruitful, operative and scientifically manageable. Getting to such an answer, we may make a distinction between *embodied learning* and *conscious learning*. The difference is between not acquiring and acquiring a belief where beliefs are extrinsic properties of the brain. Being a conscious experience, which in and by itself is an extrinsic property of the brain, puts the organism into a state where it *believes* what the sensory state presents to it. Such a belief is possible if and only if the sensory presentations possess qualitative as well as conceptual features. Indeed, a belief is a propositional attitude towards the content of a sensory state; thus, for a sensory state to produce such a belief, it must contain qualitative and conceptual elements.

Embodied learning happens *without the organism needing to pay attention to what it learns in order to use this learning*. In other words, the learning organism does not need to form any belief about its sensory states for these states to cause an appropriate behavior. In contrast, conscious learning happens *whenever an organism needs to be attentive to what it learns in order for it to use this learning*. In such cases, an organism has to believe its sensory presentation so that that it can use the informational content of the sensory input to conduct a life-preserving behavior.

A bird gets to know its habitat though a constant updating of its memory by movement and perception. This is a kind of embodied learning. The bird has a practical non-propositional knowledge of its environment brought to it only by movement and visual sensation. The physical behavior of the bird enforces its sensory presentation (mental map) of its habitat. However, the bird orientates itself according to such a sensory presentation even though it is unable to form beliefs about the objects presented. Thereby I am not saying that birds cannot form

beliefs. All I am suggesting is that its environmental orientation probably does not belong to the category of belief-generating sensations. However, the pair of back birds in my garden may present cats and dogs differently and therefore be able to have different beliefs concerning their experience of one or the other.

Opposite to embodied learning, we have conscious learning in those cases in which an organism is aware of the content of its sensory presentation because it believes or holds another propositional attitude towards this. However, being aware of these elements requires that mechanisms of its nervous system enable the organism to individuate and conceptualize sensory information. There is nothing to believe if the information processing process in an organism does not allow it to notice the differences and similarities among things. An organism that is able to do conscious learning has to be neuronally adapted to distinguish and assemble conspicuous signals from its environment and to identify them again individually as well as typologically. It has to generate conceptual presentation by abstraction to practice conscious learning, and the knowledge it acquires would propositional knowledge.

Beliefs are prerequisites for thinking. It is easy to see why thinking may help organisms to survive. Producing beliefs is a way of identify the object of one's experiences and adjust once behavior accordingly. The black bird sets off alarms when it sees a cat sneaking around, while it does not take much note of a dog patrolling the garden. However, most animals, and human beings as well, are not aware of their actual beliefs and therefore are not conscious of what they are conscious about. To be conscious of one's own beliefs demands that one is aware of having them. Apparently, only a small fragment of non-human animals are self-aware of their own thinking, which means that their brain is not adapted to making a presentation of beliefs about their sensory experiences. When it comes human beings, they have the ability to present their own thinking, although we should also realize that much of this thinking is out of reach.

Conclusion

This paper proposes a distinction between intrinsic and extrinsic properties that we may attribute to the brain. In general, intrinsic properties are those that occupy neuroscientist most, whereas extrinsic properties are what cognitive scientists commonly reflect on. Hence, my suggestion is that cognitive neuroscience can overcome the various hard problems that have tormented the philosophy of mind for so long. In biological systems, intrinsic properties of their brain are adapted to present information that is brought to them by sense organs and/or information coming from other parts of their body. Such a presentation is possible as an extrinsic property of the brain. Extrinsic properties characterize all physical systems whose intrinsic states change under the influence of causal interactions with something outside themselves. It is the capacity of presenting the surrounding world together with the actual informational input that define mentality. And if such a sensory presentation can cause beliefs about the world in an organism, then I argued, this creature is conscious about the world. That is primary consciousness. Secondary consciousness comes with the brain's ability of actually presenting its own internal state. This ability has adapted us to fine-tune our thinking to the advancement of science and technology.

Literature

Ardiel, E.L. & C.H. Rankin (2010). "An elegant mind: Learning and memory in Caenorhabditis elegans." *Learning and Memory* 17: 191-201.

Armstrong, D.M. (1970). "The Nature of Mind." In C.V. Borst (ed.) *The Brain-Mind Identity Theory*, 67-79. London: Palgrave Macmillan.

Carruthers, P. (2017). The Illusion of conscious thoughts. *Journal of Consciousness Studies* 24(9-10): 228-252.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Churchland, P.M. & P.S. Churchland (1981). "Functionalism, Qualia, and Intentionality." *Philosophical Topics* 12(1): 121-132.

Dennett, D. (1997). "Quining Qualia." In Ned Block, Owen Flanagan and Güven Güzeldere (eds.) *The Nature of Consciousness: Philosophical Debates*, Cambridge, MA: The MIT Press.

Dretske, F. (1995). *Naturalizing the Mind.* Cambridge, Mass: MIT Press.

Faye, J. (2019). *How Matter Becomes Conscious. A Naturalistic Theory of the Mind*. London: Palgrave-Macmillan.

McGinn, C. (1989). "Can We Solve the Mind-body Problem?" *Mind* 98: 349-366.

Nagel, Th. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, 83(4): 435-450.

Rosenthal, D.M. (2002). "Explaining Consciousness." In David Chalmers (ed.) *Philosophy of Mind. Classical and contemporary Readings*, 406-421. Oxford: Oxford University Press.