Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-020-00591-y OPEN

Two conceptions of consciousness and why only the neo-Aristotelian one enables us to construct evolutionary explanations

Harry Smit[1]✉ & Peter Hacker[2]

Descartes separated the physical from the mental realm and presupposed a causal relation between conscious experience and neural processes. He denominated conscious experiences 'thoughts' and held them to be indubitable. However, the question of how we can bridge the gap between subjective experience and neural activity remained unanswered, and attempts to integrate the Cartesian conception with evolutionary theory has not resulted in explanations and testable hypotheses. It is argued that the alternative neo-Aristotelian conception of the mind as the capacities of intellect and will resolves these problems. We discuss how the neo-Aristotelian conception, extended with the notion that organisms are open thermodynamic systems that have acquired heredity, can be integrated with evolutionary theory, and elaborate how we can explain four different forms of consciousness in evolutionary terms.

[1] Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, P.O. Box 6166200 MD Maastricht, The Netherlands. [2] St John's College, Oxford OX1 3JP, UK. ✉email: h.smit@maastrichtuniversity.nl

Introduction

This paper discusses essentials differences between two conceptions of mind, namely the Cartesian and neo-Aristotelian conception. It argues that only the neo-Aristotelian conception is coherent and can therefore be integrated with evolutionary theory (resulting in explanations and, after these have been further elaborated, in testable hypotheses). We shall argue that the alternative Cartesian conception is for logical or conceptual reasons incoherent, clarifying why attempts to integrate it into evolutionary theory have not resulted in evolutionary explanations.

Descartes distinguished the physical from the mental realm. The mind is according to Descartes a separate immaterial substance (its essence is thinking) causally connected to the brain. Descartes' dualism is related to the distinction between humans and the other animals: only humans possess a mind, whereas the other animals (and plants) are mere machines. Aristotle, the biologist of the mind, argued that what is distinctive about the human *psuchē* is that it incorporates not only the vegetative powers of growth, nutrition and reproduction, and the sensitive powers of perception, desire and motion, but also the rational capacities of intellect and will. The *psuchē* is not, in contrast to what Descartes believed, causally connected to the body, but is characterized as the 'form' of the living body. Aristotle characterized human beings as distinct from animals and plants by reference to what they can do (i.e. their capacities expressed in behaviour, including linguistic behaviour). We shall argue that, for understanding the evolution of consciousness, we must extend Aristotle's conception of the vegetative, sensitive and rational *psuchē* with the insight that living organisms are open thermodynamic systems that have acquired heredity. As the result of evolutionary transitions complex thermodynamic systems evolved, e.g. cells with cell-organelles and later multicellular organisms with organs. We explain how four forms of consciousness evolved by integrating evolutionary theory with Aristotle's discussions of the capacities the different organisms possess.

What is the mind?

We start by briefly discussing investigations of the concept of mind, for this clarifies why there are two opposing conceptions. The question 'What is the mind?' was answered by philosophers by giving a definition. This is understandable: 'the mind', being a nominal, invites the question what kind of entity it is. Descartes defined the mind as an immaterial substance the essence of which is thinking. Cognitive neuroscientists, though critical of Descartes' dualism, still conceive of the mind as a kind of thing: they identify the mind with the material brain. Neo-Aristotelians, by contrast, argue that the question 'what sort of entity is the mind?' is as misleading as the question 'What sort of entity is a number?'. We should according to them start with a different question: what must be true of a living creature for it to have a mind? Answering this question reveals that the mind is mainly associated with intellectual faculties (Rundle, 1997, chapter 2). For example persons are said to have a powerful, agile, or subtle mind if they are skilful, quick and ingenious at problem solving. If someone holds something in mind, he retains it in memory; if she knows her mind she has formed an opinion; and if he makes up his mind he decides. A person is of sound mind if he retains his rational faculties, and out of mind if he thinks, proposes or does things that are irrational. If persons have lost their mind they are bereft of their rational faculties. This characterization of the mind does not imply that it is an entity. When we say that someone 'changed his mind', that he has 'a dirty mind', or that he has 'turned his mind to a certain question', we do not mean that there is a single entity that has changed, is dirty, or has been turned. We then speak of a human being (not of a mind) *of whom* we say different things. To have a mind is not to be in possession of a kind of entity, but to possess a distinctive range of powers, namely the rational powers of a language-using creature. Our rational powers endow us with a horizon of possibilities of action, thought and feeling. They enable us to reason and to act for reasons.

Such conceptual investigations teach us that the mind is not an entity or thing of any kind, and reveals an association of the concept of mind with rational faculties. The latter is the reason why the alternative to the Cartesian conception is called a neo-Aristotelian one. For what Aristotle called the rational *psuchē* coincides roughly with what must be true of a creature to have a mind (to possess and exercise intellectual powers that presuppose mastery of a language).[1] The neo-Aristotelian characterization has not only the advantage that it corrects the Cartesian misconception of the mind as an immaterial substance, but also the misguided association made by Descartes between consciousness and life. While Descartes sharply distinguished the mind from the body and brain (in his view, life is linked to consciousness only in humans animal life being wholly mechanical), neo-Aristotelians, following Aristotle's demarcation of the animate from the inanimate, argue that plants possess powers too, namely vegetative powers, and animals in addition sensitive powers (the vegetative and sensitive *psuchē* were held by Descartes to be explicable in mechanical terms).

The mystery of how the brain codes consciousness

Descartes distinguished the mental realm from the physical realm (for a discussion of his ideas, see Ben-Yami, 2015). He argued that only humans have a mind, that the mind is a separate immaterial substance, and that the mind is causally connected to (the pineal gland of) the brain. He assumed that we can understand the physical realm in terms of the laws of mechanics. Since animals lack a mind (they are mere automata), Descartes argued that we can understand animal behaviour in terms of physics (he believed that explanations in biology are reducible to the laws of physics). The mental domain, by contrast with the physical domain, does not consist of matter (the essence of which is extension). Descartes characterized the psychological realm in terms of consciousness: the awareness of thoughts or experiences within us. He redefined thought to incorporate 'everything which we are conscious of as happening within us, insofar as we are consciousness of it' (*Principles of Philosophy*, I, 9). Accordingly, perceptual experience, imagination, cogitation, affection and volition constitute the domain of consciousness. They are different forms of thinking. We cannot think thoughts (have experiences) without being conscious of so doing, and cannot be conscious of thoughts without knowing that we have them.

Descartes' dualism raises the difficult question of whether conscious experiences are explicable in terms of brain processes. How can events in the objective, physical world (for example neural events) give rise to events in the subjective, psychological world (conscious experiences)? The reason why this question seems hard to answer is that consciousness does not belong to the world studied by the physical sciences. On the other hand: we cannot deny that we have conscious experiences. We are conscious of all the thoughts and experiences that pass through our minds. As Tononi and Koch (2015, p. 2) put it: 'That I am conscious, here and now, is the one fact I am absolutely certain of —all the rest is conjecture'. How, then, can we reconcile consciousness or subjective experience with the Cartesian conception of objective reality as purely mechanical?

These problems return if we ask, as cognitive neuroscientists (e.g. Crick, 1995; Dehaene, 2014; Gazzaniga, 2018) and (evolutionary) psychologists (e.g. Ginsburg and Jablonka, 2019; Tooby and Cosmides, 1992) do, how the brain codes consciousness or thinking. We presuppose then that the brain is the organ of consciousness, for what other organ can fulfil this role? However, it is easy to see why this presupposition immediately raises difficult questions: while for example the eyes are the organs of perception, it is unclear in what sense the brain is the organ of consciousness. We see with our eyes, but are not conscious (e.g. awake) *with* our brain (although we cannot be awake without a brain either). We also do not become conscious or aware of something (e.g. the patter of rain drops on a window pane while we are reading) with our brain, and cannot write our name with our brain, for the brain is not an organ the movement of which we can voluntarily control. We cannot *use* our brain as we can use our senses or external organs like our hands. These observations make the question of how consciousness emerges during ontogenesis (the development of an individual) and emerged during the course of evolution seem not only problematic, but also mysterious. The classical example of how we perceive a red apple may help us to understand the roots of our puzzlement.

Suppose light is reflected off a red apple into our eyes. We then see a red apple, say so, and perhaps reach out to take it. We can also describe the seeing of the red apple with the aid of knowledge from physics, chemistry and neuroscience. Light of a certain wavelength, reflected off the apple, excites the rods and cones of the retinae. It is absorbed by rhodopsin and photopsin (proteins) present in the membrane of respectively rod and cone cells, resulting in the phototransduction of light into an action potential. This culminates in the conduction of impulses along the optic nerve to the thalamus and the (striate) cortex.[2] Notice that, so far, there are no mysteries: we have described the visual experience in terms of (1) my behaviour (how I respond when I see a red apple and what I say when I see the fruit) and (2) the physical and biochemical processes occurring in my eye and brain when I see the red apple. However, problems arise if we ask the Cartesian question of how objective neural events in our brain are transformed into a subjective or conscious *experience* in my mind or brain. How does a subjective experience emerge out of a complex arrangement of material particles? And it raises the question at what point on the evolutionary scale the conscious experience of red suddenly emerged (for it seems to be an all-or-nothing phenomenon). Some cognitive neuroscientists and philosophers argue that this is the real mystery or hard problem (Nagel, 1974; Chalmers, 1996). There are several reasons why Descartes' characterization of thought leads to unresolvable problems.

First, notice that we do not have any empirical reason for believing that there are experiences in the brain. One may enjoy an experience in the garden, theatre or home; one may have a headache in the library or a backache in the fields. But there is no such thing as enjoying a performance of *Romeo and Juliet* in one's brain, let alone of suffering a backache in one's head. In a similar vein: when we see a red apple, we do not see an image of the apple in the brain (and neuroscientists, when they study the brain, do not find these images either).

Secondly, there are no 'inner eyes' with which we know how things are subjectively with us. Descartes' successors (Locke, Hume, James and others) suggested that we acquire such knowledge by the exercise of a cognitive faculty. This faculty was thought to be a form of perception or inner sense, leading to the idea that we know the content of experience by apperception or introspection. However, this philosophical notion of consciousness as apperception is a misconception. There is no such thing as an inner sense with which we or the brain can perceive experiences. We do not perceive experiences (or images); we *have* them. And 'looking into' our mind does not provide us with *evidence* either. When I see for example a red rose and I say so, I introduce a report on public visibilia ('I see …') or justify such a report by citing what I saw. I need no evidence for its presence: it is evident to my senses (I see it with my own eyes!). The petals on the ground may be evidence that there was a rose in the garden, just as crumbs on the table may be evidence that there was bread on the table. But my seeing the rose or the bread is not evidence for me.

So having 'subjective experiences' is not a form of perception and does not constitute *evidential grounds* for my knowledge of the world around me. Neo-Aristotelians argue that Cartesians mistakenly believe that we are acquainted with subjective experiences in a similar way as we are acquainted with material objects. Notice that Descartes' *redefinition* of thought in terms of consciousness is the culprit here: according to Descartes, there can be nothing in the mind that is not illuminated by the light of consciousness. But this is incorrect. We are neither conscious nor unconscious of everything we perceive. Unlike observing, looking at, scrutinizing, or listening, being perceptually conscious of something is a form of cognitive receptivity. One cannot voluntarily or intentionally become or be perceptually conscious of something, and one cannot order someone to be conscious of something. One can be conscious only of such objects of perception that catch and hold our attention. Similarly, when we feel a pain we become and are conscious of the sensation, but, as we have seen above, we do not have *access* to the content of our experience (to be conscious of something belongs to category of capacity, not to the category of introspection). If I become conscious or aware of a pain, then my attention is caught and held by my pain and its severity. Of course, we can *express* our perceptions and sensations in natural behaviour and with the aid of mental concepts. Yet neo-Aristotelians emphasize that these expressions have no *foundation* in subjective experiences, as Descartes thought (this is further elaborated in seventh section). For example 'It looks … to me' is not to *apperceive* and to *describe* a subjective experience. The sentence one utters is a linguistic *expression* (not a description of the content of an experience) of what one takes oneself to be observing (made possible by our discriminatory capacities and mastery of a language). How things perceptually *strike* us is expressed by saying 'It looks … to me': it is an impression consequent upon looking at something. And if one is wrong (something is not as it strikes one, for example when it is seen in better light), then one will say 'Oh, this is not as it initially appeared to me to be'.

Consciousness as the upshot of information-processing computations

Other philosophers (and cognitive neuroscientists) also noticed these problems confronting the notion of apperceiving the content of experiences and have constructed alternatives to, what they call, 'the Cartesian theatre' (Dennett, 1993). According to them, the varieties of thought or conscious experiences are accomplished by the brain. We should think of the brain as a computer, as an information-processing system. Yet what is noteworthy in the context of this paper is that they still adhere to the Cartesian conception of consciousness (see Bennett and Hacker, 2003). According to them, being consciousness means not 'being conscious of something' (as it is ordinarily and by neo-Aristotelians understood), but 'having an experience of something'. What Descartes called 'thoughts' is roughly coextensive with what they call 'experiences'. These experiences are now conceived as the upshot of 'information-processing computations' in the brain operating on sensory inputs (see e.g. Dehaene, 2014;

Dennett, 1993; Tononi, 2005). The computations result at first in non-conscious 'representations', but when information is further 'integrated' at higher levels (in the 'workspace of the mind'), they may result in conscious experiences. These are known by introspection, which is conceived as a form of 'meta-cognition' (it is assumed that there is a 'self-monitoring system' present in the brain) that allegedly makes use of internal representations. In order to explain why these ideas do not resolve problems, we discuss the example of how a driver becomes conscious of the flashing fuel tank light.

We first elaborate the example in terms of the ordinary or neo-Aristotelian conception of consciousness. According to this conception the object of consciousness is not a mental representation, as cognitive neuroscientists believe, but the flashing light (the driver is conscious *of something*). The driver becomes conscious of the flashlight when it catches and holds his attention. Dehaene, by contrast, characterizes the situation as a *relation* between a cognitive system (in the brain) and a specific object of thought, namely the mental representation of 'the fuel tank light' (Deheaene et al., 2017, p. 486), and argues that consciousness 'arises' in the brain. We mention five reasons why this is misguided. First, 'The driver is conscious of the fuel tank light' does *not* refer to a *relation*, but to the driver. Second, it is the driver who becomes conscious of something, not a cognitive system in the brain. Third, when the driver becomes conscious, the object of his consciousness is not a mental representation but the flashing light. Fourth, consciousness does not arise in the brain (as Dehaene argues); it 'arises' in an animal that has the appropriate kind of brain. Humans *and* the other animals *become conscious* (not cognitive systems in their brain), are subsequently conscious of things, and are thereby in receipt of knowledge. Fifth, when the driver is conscious of the light (not of an experience he has in his brain), be becomes aware of the need for fuel. He may subsequently *think* that he needs to stop at the next petrol station (notice that the object of *that* thought is not a light on the dashboard).

One might object that the use of the notions of representation and information-processing are essential for understanding the explanation of cognitive neuroscientists. However, these notions do not clarify issues either. First, the notion of representation is left unexplained. Dehaene believes that processed information becomes available as a representation when it catches and holds someone's attention. But increased firing of neurons (as the result of sensory input) does not represent what causes it, either in a pictorial or in a symbolic sense. The firings are indicative of something, but do not represent what they indicate, just as smoke indicates fire but does not represent it in the sense in which a painting or description of a fire represents it. Hence when we say that there is a 'neural correlate' between a perceived phenomenon and increased firing of neurons, it is ill-advised to say that it results in a mental representation. Second, the idea that *conscious experiences* arise in the brain as the result of information-processing does not clarify issues either. The reason is that cognitive neuroscientists do not clearly distinguish information in the mathematical or information-theoretic sense of Shannon and Weaver (Shannon and Weaver, 1963; Weaver, 1949) and information in the ordinary sense. I obtain information from perception (or from what I read or hear from others), but my brain cannot obtain information in this ordinary sense (although it can be said to transmit information in Shannon's sense, but that is *not* ordinary information). Let us explain.

Cognitive neuroscientists like Dehaene and Tononi presuppose that there is sensory input (e.g. sound waves agitating our ear drums or photons impinging on our retinae) which is subsequently processed by our brain, resulting in cognitive output, namely meaningful statements or assertions about the external world. Notice that photons and sound waves are not 'unprocessed information'. When I tell someone that p (e.g. 'It is raining in Amsterdam'), then I provide him or her with information, but the stream of photons or sound waves is not information in this sense. It is important not to conflate here the language of physics and ordinary language. Suppose that I listen to someone and respond by saying something. Physicists describe the input in terms of sound waves and, since the phrase 'meaningful assertions' does not belong to physics, they also characterize the output (my response) by sound waves. But since I am providing someone with information (I am using a language), we can also describe the input and output in terms of information (involving the meaningful use of words and sentences). A speaker may tell someone something in order to inform the hearer about what he has observed, and by listening to his words and sentences the hearer acquires information. In- and output are then characterized in terms of intelligible verbal assertions. We must not conflate these two possibilities.

The computer-metaphor does not clarify issues either. Humans acquire information; computers do not. For in the information-theoretic sense (see below) in which computers process information, neither understanding nor knowledge is involved. Computers do not understand what they are doing; they simply causally compute without knowing how they do it. What computers do is a causal operation, not a normative, rule-following operation (which can correct or incorrect, etc.). Computers are designed by us to produce the results of computation but do not compute anything, just as slide rules produce the results of computation without actually computing anything.

Though brains can be said to process Shannon-information, the information they then process is independent of meaning. Shannon asked: if 'x' is a discrete random variable, to what extent is 'information' received when a particular value of 'x' is detected following transmission in an electronic system? Shannon postulated that if a highly probable value of 'x' is detected, then little 'information' is gained. The opposite is the case with a low probable value of 'x'. The quantity of 'information' $h(x)$ can be obtained from the probability distribution $p(x)$ of x. This is Shannon-information, defined by $h(x) = -\log_2 p(x)$. As an illustration: tossing a coin and obtaining heads corresponds to $\log_2(2) = 1$ bit of information, because there are just two alternatives; throwing a dice yields $\log_2(6) = 2.59$ bits of information, because there are six equally likely possibilities.

Based on Shannon's ideas, Tononi (2005, pp. 111–112) argues that, 'every time we experience a particular conscious state out of such a huge repertoire of possible conscious states, we gain access to a correspondingly large amount of information'. The occurrence of a particular conscious state is extraordinarily informative because of the very large number of alternative conscious states that it rules out. Tononi compares the processing of information by the brain to the way a photodiode processes information. A photodiode changes its electrical resistance depending on the illumination. If it detects complete darkness it generates a minimal amount of information, since it can only discriminate between darkness and light. When we consciously detect complete darkness, however, we perform a discrimination that is immensely more informative: we are not just ruling out light, but an extraordinary number of other possible states of affairs, including every possible frame of every possible movie, every possible sound, and every possible combination of them. Hence whenever we experience a given conscious state, 'each of us is gaining access to an extraordinarily large amount of information' (Tononi, 2005, p. 11). This is highly problematic. Suppose I see an owl in my garden. I may share this observation with others by telling them that 'I saw an owl in my garden'. But neither I nor my brain have access to the

information that the owl was one of thousands of other bird species, that it was sitting on the branch of an oak and not on the branch of thousands of other trees, and so on and forth. I gain no access to any such things, and my brain has no access to this information either. The proposition 'p' describes the state of affairs that p ('There is an owl in the garden'). It can be true and can be false. If p is true (this possibility is actualized), then it excludes the alternative possibility (namely not-p), but it does not exclude other possibilities (not-q, not-r, etc.). Tononi mistakenly believes that if I say p, I exclude not only not-p, but also not-q, not-r, and so on. Yet not-p does signify all the possible states of affaires in the universe other than p, otherwise not-not-p would not be p. Of course, one can argue that there is Shannon information (related to impulses moving along nerves) processed by the brain. But Shannon information has nothing to do with ordinary information (which we obtain by using our faculty of sight, or from hearing or reading). We gain no information of neural impulses moving along the optic nerve (though the neuroscientist may derive information from it), while we gain information from seeing something, e.g. an owl.

Summarizing: neo-Aristotelians argue that ordinary information can and should sharply be distinguished from Shannon information. As long as we clearly distinguish Shannon and ordinary information, we can answer how the brain 'processes' Shannon information and how humans convey ordinary information when they chat and so on. Yet if one starts with the assumption of modern cognitive psychology and cognitive neuroscience that the brain is a sort of computer, one is easily seduced to conflate ordinary information and Shannon-information. It leads to the misguided idea that we have conscious access to experiences as the result of information-processing computations in the brain.

We can conclude that it is the Cartesian conception of consciousness that leads to the misconception that the object of consciousness is (a variant of) a Cartesian thought. Becoming conscious of something (an object, a sensation) is not gaining knowledge as the result of apperception, nor is it the upshot of information-processing computations. We have seen that to become and to be conscious of an object is a form of *receptive knowledge*: our attention is caught by something on the *periphery of our visual field*, for example the flashing light indicating need of petrol. We shall later explain why, as the result of natural selection, animals became sensitive to, for example, moving objects in the periphery of their visual field. We can also conclude that, when we investigate conscious experiences, we do not encounter problems as long as we do not redefine our ordinary conception of consciousness in Cartesian terms. If we start with the ordinary conception, we can ask *how it is possible* for us to become conscious and to express it in natural behaviour and linguistic behaviour. In terms of evolutionary theory: how did the capacity evolve to *express* being conscious of something in natural behaviour and linguistic behaviour? If we ask these questions, we are no longer haunted by Cartesian mysteries. For we have seen that these arise if, when we re-interpret the *capacity* to express conscious experiences in Cartesian terms, we ask how a *brain state* can code a conscious experience. It is the shift from capacity to state that generates puzzlement. Moreover, while abilities, just like skills, come in *degrees* and can therefore *gradually* result in an improvement during development and evolution, it is unclear how having a conscious experience evolved.

In the following we shall explain why neo-Aristotelian monism is the conceptually sound alternative. Therefore we elaborate the neo-Aristotelian conception, how we can integrate it with evolutionary theory, and then explain how we can understand four different forms of consciousness in terms of modern evolutionary theory.

Powers and the Aristotelian conception of the *psuchē*

While Descartes distinguished two domains (dualism), Aristotle argued that there is only one (monism). Aristotle characterized human beings as distinct from other animals and plants by observing their *capacities*, i.e. what they can do. Plants have a vegetative *psuchē*, animals have in addition a sensitive *psuchē*, and humans alone possess a rational *psuchē* (the power to reason and to give reasons, a capacity that is absent in the other animals). According to Aristotle the *psuchē* informs a natural body that has life and this is visible in the exercise of capacities. And since the *psuchē* is ascribed to living organisms, the capacities that constitute the *psuchē* cease to exist when the living creature dies. The *psuchē* is not a separate, immaterial substance, as Plato (and much later, Descartes) believed, but the form of the body that has life.

Aristotle's discussion of the differences between plants, animals and humans can be integrated with evolutionary theory if we notice that organisms are not machines (in the case of humans alone united to the mind), but thermodynamic systems (Smit, 2018). Organisms are open thermodynamic systems that have a metabolism (a coordinated system of chemical reactions contributing to its maintenance, i.e. a system that imports energy to maintain order, but also to grow, reproduce, et cetera) and hereditary replication (a system of copying in which the new structure resembles the old). They evolved from simple organisms like a cell to complex ones like plants and animals. Like Aristotle, we can characterize these different forms of life by reference to their powers, which can only be studied by investigating the behaviour of organisms, for activities and actions are logically prior to potentialities. Aristotle characterized the lower forms of life as the nutritive or vegetative *psuchē*. Plants have the powers of metabolism, growth and reproduction. One can of course argue that unicellular creatures have these features too, but notice that there is an essential difference: in multicellular plants organs evolved which execute different functions (for example roots for absorbing water and nutrients). The life of animals is, over and above the vegetative *psuchē*, characterized by the sensitive *psuchē*, the primary form of which is touch and taste. And when sensitivity evolved, the capacity for pleasure and pain evolved too resulting in appetite. This enabled the primitive organisms living in the sea and later on land to nourish themselves and to distinguish what is beneficial from what is detrimental to them. As an extension of these primitive forms of sensitivity, the senses of sight, hearing and smell evolved together with the powers of locomotion. Whereas plants only have needs (for example they need water), animals also have desires and aversions and act for the sake of a goal. Humans, as language-using creatures, possess on top of the sensitive powers a rational *psuchē*. The power of reason is the ability to reason, i.e. apprehend the transition from premises to conclusion as warranted. Any creature that can reason is also sensitive to reasons, reasons for acting or refraining, reasons for thinking something to be so, and reasons for feeling. And any creature that can reason and can do things for reasons can also answer the question 'Why?'. So they are answerable for their deeds.

This elaboration of the neo-Aristotelian alternative enables us to explain why it is not confronted by the problem of how conscious experiences evolve out of matter. There is a simple reason: neo-Aristotelians do not ask the Cartesian question of how conscious experiences evolve out of matter (or physical or chemical systems), but ask how our susceptibility to have our attention caught and held by items on the periphery of our visual field evolved in our ancestors, who, like the other animals, were self-moving organisms possessing senses. The rational *psuchē* evolved as an extension of the sensitive *psuchē*. There are two main reasons why Cartesians and neo-Aristotelians ask different

questions. First, Descartes held the vegetative and sensitive *psuchē* to be explicable in mechanical terms, clarifying why Cartesians do not ask how the human mind gradually evolved as the result of modifications of several existing capacities or powers present in our ancestors (which they shared with the other animals). Second, we have seen that Descartes, in contrast to neo-Aristotelians, did not hold the mind to be a distinctive set of powers of an organism. He conceived the mind to be a separate, immaterial substance. Descartes found it necessary to recognize an *inner world* (conscious experiences) distinct from the outer world. This inner world was by his successors thought to consist of mental objects (ideas, images), events (a sudden recollection), states (of joy or sorrow), and processes (thinking, imagining). We have seen that this conception of the mind leads to unresolvable problems. We shall now explain the main reason.

The general problem with the Cartesian conception is that the idea of an immaterial, thinking substance makes no sense, since *there are no criteria of identity* for such a thing. We can identify a substance (e.g. a material object, plant or animal) in the external world, but what is meant by an immaterial substance? How can we identify it? Notice that this is not an epistemic but a logical problem, for nothing has been laid down to determine what would count as one such substance as opposed to a thousand. As Kant once put it: how can we distinguish one immaterial substance thinking all my thoughts from a thousand such immaterial substances thinking the same thoughts? Furthermore, if we cannot identify the immaterial, thinking substance, then it also makes no sense to say that there is a *relation* (or interaction) between an immaterial mind and the material brain or body. Consequently, there are not two worlds (that interact). We have to abandon Cartesian dualism because it is illusory. The 'mental' is not *inside* anything, as if there were an internal world populated with mental objects, processes, states, and acts, and the physical is not *outside* the mind, since the mind is not a space. Thus we can conclude that the notion of an immaterial, thinking substance is incoherent, clarifying why scientists and philosophers after Descartes were puzzled by, but could not resolve, the problem of why and how conscious experiences, as they were conceived by Descartes, *relate to* or *interact with* the material brain.

The reason why the neo-Aristotelian conception is not confronted by unresolvable problems it that it (coherently) characterizes the mind in terms of capacities or powers, and powers can gradually evolve during the course of evolution. These powers do not belong to a separate substance for they are not material or immaterial things. We can ask how the *psuchē* relates to a living being. Is it a part of the living organism just as organs, such as the eye, are a part of the organism? No. Eyes are parts of organisms and we cannot see without eyes. But eyesight is not a part of the eye. Without eyes, there is no eyesight, but eyesight does not make the eyes see, just as the power to cut wood does not make the axe cut. Animals, i.e. self-moving organisms with senses, see with their eyes (just as legs are the organs of walking, and the mouth the organ of speaking). Recall here that the human being as a whole sees, not his or her eyes. So too it is the human being that exercises his or her rational powers, who reasons, infers, comes to conclusions and acts for reasons, not the rational *psuchē*. To say that the *psuchē* reasons or deliberates is, as Aristotle (*De Anima*, 408b12–408b15) put it, like saying that 'the *psuchē* weaves or builds. Surely it is better not to say that the *psuchē* pities, learns or thinks, but that the man does these things with his *psuchē*'. The *psuchē*, in contrast to an organ or a cell, is not a part of a living being. We can do things with our *psuchē* in the sense in which we do things with our talents. Humans have abilities and capacities, liabilities and susceptibilities. But while an axe stands in a relation to its owner (i.e. is possessed by someone), having rational powers is not a relation of ownership between a person and a set of powers. That is also the reason why it is absurd to ask whether the body and *psuchē* are one or two. The *psuchē* is not a part of a body for the *psuchē*, being the form of living things, is inseparable from the living body. It is the principle of life characteristic of kinds of multicellular living beings, for its distinctive powers are what make a living being with organs the kind of being it is.

Four forms of consciousness and why they evolved

Neo-Aristotelians argue that we should characterize the mind in terms of powers. Evolutionary theory teaches us that powers characterizing the different forms of organisms evolved during the course of evolution. For understanding how we can investigate these powers in terms of the (evolution of) brain and behaviour, it is important to notice that there have been several evolutionary transitions (Maynard Smith and Szathmáry, 1995). The transition from the symbiotic (or eukaryotic) unicell to multicellular organisms is relevant for our discussion of consciousness.

The first multicellular organism evolved out of a cluster of cells that were genetically identical (they were all derived from a single cell). Hamilton's inclusive fitness theory (Hamilton, 1964) clarifies why, because $r = 1$ (r is the coefficient of relatedness), cooperative interactions between the cells could easily evolve. For example cells started to exchange molecules between them and later extended these cooperative interactions by producing ligands and receptors resulting in chemical communication. One cell could then signal to another cell and alter its behaviour. Another innovation was cell differentiation as the result of new mechanisms of gene regulation. When the differentiated cell state was also maintained during cell division, specialized cells evolved resulting in organs (e.g. the brain and sense organs) that could execute different functions in the organism. Two points are noteworthy. First, innovations at lower levels (e.g. cell differentiation, migration, and so on) facilitated the evolution of more complex behaviour of the whole organism (Kirschner and Gerhart, 2005). Secondly, as the result of the transition to multicellularity natural selection started to act at the level of the whole behaving organism (Michod, 2007). Hence as the result of the transition to multicellularity more and more complex behaviours emerged (including, in the case of the human species, linguistic behaviour, e.g. the ability to use linguistic expressions) made possible by evolution at lower levels (cells and organs). This clarifies why we do not need to postulate a separate substance for understanding the emergence of new capacities at the level of the whole organism. The *psuchē* did not evolve as a separate substance, nor did it evolve as an emergent property of the brain. For example as the result of natural selection humans can see a flower in the garden and can say what colour it has, but this ability is not located in a cell of the striate cortex and neurons are not doing the seeing (as Crick, 1995, p. 104, mistakenly suggests). If someone has poor eyesight, we observe that he (not his eyes or his brain) has problems in finding his way around (he bumps into things or falls over things and cannot find things by looking, etc.). Seeing is a predicate applicable to the whole, behaving human being as a whole, using his eyes, not to parts of a human being (eyes or the brain). Of course, evolution may also lead to emergent properties of the brain. But the behaviour of an organism is not an emergent or supervenient property of the brain. That we see with eyes, mean something, or can use our hands to point and to write our name is not a property of the brain. Notice that this insight provides us with evolutionary elaboration of *a principle of mereology* (Smit and Hacker, 2014). This conceptual principle clarifies why it makes sense to attribute psychological predicates

to the behaving organism as a whole, but not to its parts (genes, cells or an organ like the brain).

When, in the emergence of different life forms, did consciousness evolve? Light-sensitive, sound-sensitive and smell-sensitive cells and organs gradually evolved in multicellular organisms (although primitive forms are present in unicellular creatures). At the early stages of evolution there is no question of consciousness, for there were only varying degrees of sensitivity to light, sound and smell. For example if the flexibility in light-sensitivity depends only on the states of the protein rhodopsin (see footnote 1), then the flexible response is no constitutive evidence for saying that consciousness or anything has emerged. But there has been a gradual extension of light sensitivity as the result of the duplication and modification of opsin genes (resulting in for example colour sensitivity) and as the result of the evolution of senses and the expansion of the brain (Gehring, 2014). When eyes, ears and noses evolved (and when the complexity of their brains increased), organisms possessing these sense-faculties were taking the early steps on the route to creatures that can be said to have their attention caught by what they perceive.[3] They then gradually acquired the flexible capacities of perceptual transitive consciousness. This observation raises an evolutionary question: what, then, were the selective advantages of consciousness? We discuss briefly four examples: intransitive consciousness, perceptual consciousness, somatic consciousness and self-consciousness.

(1) *Intransitive consciousness*, we have seen, is characteristic of sentient creatures, i.e. animals that have the powers of sensation and perception and are susceptible to pleasure and pain. Animals have a diurnal cycle of sleeping and waking and, hence, enjoy intransitive consciousness. We can use here evolutionary explanations of diurnal cycles clarifying why intransitive consciousness evolved. Notice that we distinguish being awake and asleep from being conscious and unconscious. This is related to the contexts in which we observe the responses related to the use of these verbs. 'Being unconscious' and 'recovering consciousness' are typically used in the hospital, while asking whether someone is asleep or has woken up belongs at home. Why do we differentiate between these cases? This is related to responsiveness: behavioural observations show that we are far more responsive during sleep than during periods of unconsciousness (it is not difficult to come up here with an evolutionary and biomedical explanation for this difference). Of course, there are interesting borderline cases between intransitive unconsciousness and consciousness: in rare cases humans are barely conscious, sleep–walking, groggy, or, in technical terms, in hypnotic trance, in a fugue state, et cetera. In line with what has been elaborated above it is important to notice that (un) consciousness is a state of a creature, but not a mental state. Consciousness, we have seen, is a condition for being in any occurrent mental state. Thus a state of intense concentration is a state of consciousness we are *in* (for example when we are doing an exam) while we are conscious. This occurrent mental state is opposed to a dispositional mental state, e.g. being in a depression.

(2) *Transitive perceptual consciousness* concerns (among other things) the passive power to have one's attention caught by things and events in the periphery of one's perceptual field. It signifies a form of coming to know or of knowing something. Other verbs we use to signify this are being aware, noticing, realizing and recognizing. These are all factive verbs of cognitive receptivity. The point to notice is that the verbs used for cognitive receptivity are not *achievement* or success verbs but *result* verbs. In contrast to verbs such as 'detect', 'discern', 'discover' or 'find out', they do not signify achievement of knowledge (Hacker, 2012, 2013; White, 1964). They indicate the reception of knowledge: we become conscious or aware of something, notice, recognize or realize something, but this is not a matter of achieving knowledge as the result of our endeavour, for these are not acts or activities that we may perform voluntarily, intentionally, deliberately or on purpose. We cannot ask 'How are you conscious of…?'. We ask 'How do you know …?' but 'What made you conscious of …?'. For while there are sources of knowledge (perception, evidence, reason, etc.), there are no sources (as opposed to causes) of being conscious of something. This also clarifies why we cannot be ordered or forbidden to become or be conscious of something. It also clarifies that consciousness of something is neither an exercise nor a skill (when we are conscious of something, we are not an expert or authority in a given domain of knowledge). We cannot become skilful at becoming conscious of things, although we can be trained in greater receptivity and to be more sensitive to certain things. Since the other animals are sentient beings too, they also have this passive power. How did it evolve?

Take the example of a gazelle grazing on the savannah. When it is searching its surroundings for predators, the gazelle can attend to the colour and contours of a predator. When we explain to another person what we see when a gazelle is paying attention to a predator, we describe what the gazelle is doing: it is for example frequently looking in the direction of the predator (observing, watching, scrutinizing). The point to notice is that although looking with and without attention may not differ, what the gazelle subsequently does if it looked with attention may differ greatly (for example when the predator moves one step in its direction it may immediately start fleeing). Hence if the predator *captures* its attention, the gazelle is conscious of what holds its attention. We may notice this as observers since attention may have a manner: the attention paid to the predator may be intense or careful.

Now imagine a gazelle grazing on the savannah and assume that the animal has not yet noticed a predator. Since attending to something may be intermittent, the gazelle will first pay attention to what it is eating but will, after a while, forage mainly inattentively or mechanically. Suppose that the gazelle, engaged in (goal-directed) foraging behaviour, notices a moving creature in the periphery of its perceptual field. It will then become conscious of the presence of the creature (it realizes that it is a predator). Perceptual consciousness is then a matter of peripheral attention: the gazelle becomes conscious of impending danger resulting in the activation of the fight-flight system. The advantages are clear: an animal that lacked such a power would have a selective disadvantage since it would have higher risks of being caught by a predator. It is easy to see that we can extend this story by discussing adaptations on the side of the predator, clarifying why the 'seek and hide'-arms race between predator and prey has resulted in improved cognitive receptivity of the prey (and better fleeing skills) and improved hunting skills of the predator.[4]

(3) What are the evolutionary advantages of *transitive somatic consciousness*? Recall it is the foot that hurts, not the mind or the brain. As Aristotle taught us, we have a sensitive body: our head may ache, our back may tickle, our leg may itch and our fingers may hurt. The human body, we have seen, is not an insensate machine, but a living organism.

Since there is no difference between having a pain and being consciousness of a pain, the evolutionary advantage is patent. For example when I have a splinter in my finger because I have used a wooden tool, the pain I experience in my finger (not in my brain), induces me to remove the splinter before it may cause severe damage to my body (e.g. an inflammation). And when someone expresses his or her pain in behaviour, inclusive fitness theory teaches us that this form of communicative behaviour may elicit helping behaviour and enhance the inclusive fitness of the helping individual.

(4) *Reflective consciousness* (also called self-consciousness) presupposes mastery of a language. For only a language-using creature which possesses the concepts of anger, depression, excitement or love, can realize or become conscious of the fact that he or she is angry, depressed, excited or in love. Humans can engage in forms of self-reflection when trying to determine, for example, the nature of their feelings (e.g. whether they really love someone). It is 'calling up some memories, of imagined possible situations and of the feelings that one would have if …' (Wittgenstein, 2009, par. 587). Such soul-searching inquiries require imagination and judgement, but no 'inner eye', for there is no inner eye and nothing to perceive. But we can reflect on images and on what someone has done and said, et cetera, for we do have ideas in mind, thoughts do flash across our minds, and we often engage in reflective introspection. If it makes sense to speak of the evolutionary advantage of this form of transitive consciousness, then it is an aspect of the evolutionary advantage of possessing a language (see Smit, 2014, 2016), although one should keep in mind here that not every form of reflective consciousness must have an evolutionary advantage, just as we do not need to suppose that every kind of pain is of evolutionary advantage.

Self-consciousness evolves when children have acquired sophisticated linguistic powers (see Bennett and Hacker, 2003, section 12.6). This involves mastery of the use of proper names and pronouns, the various tenses, and mastery of psychological predicates, i.e. being able to apply them to one self without criteria and to others on the basis of behavioural criteria. The essential point to notice is that mastery of the use of the first-person pronoun by a child goes hand in hand with mastery of the use of other personal pronouns and person-referring expressions. For example the transition during child development from natural expressions of pain (e.g. cries) to 'Hurts!', 'It hurts!', and then to 'I have a pain' and 'I am in pain', is bound up with understanding the use of the declarative sentence'She or he has a pain/is in pain'. In learning to give verbal expression to his own pain, the child also learns to describe others as being in pain: the first-person and third-person pain-predications are two sides of one and the same linguistic coin. The possibility of groundless self-ascription of experience and thought is therefore acquired together with, and not antecedently to, the mastery of the linguistic apparatus for ascribing thought and experience to others (children do not first learn to ascribe experiences to themselves and only then to others on the basis of analogy with their own case or on the basis of his or her construction of a Theory of Mind; see Smit, 2020). And if they have mastered the ability to be transitively conscious of their own mental states and conditions, they can think and reflect on how things are with them, can act and also can become and be conscious of themselves as so acting, and can reflect on their own past and their character traits, preferences, motives and reasons for action. They have acquired the ability to be self-conscious.

I am absolutely certain that I am conscious

We have discussed how forms of consciousness, after the transition from unicellular to multicellular organisms, evolved in self-moving organisms using their senses. However, it seems that we have only touched upon an essential issue mentioned by Tononi and Koch (2015), which they attribute to Descartes, namely that my being conscious is the one fact I am absolutely certain of. They argue that we should construct explanations by taking the epistemic certainty of conscious experience into account (what they call the phenomenological properties of consciousness). We shall explain that the certainty of conscious experiences is not an epistemic certainty, as Tononi, Koch and many others mistakenly believe.

Tononi and Koch, just like Descartes, argue that our experiences, which we express by first-person propositions (e.g. 'I have a pain' or 'I feel pain'), are most directly and intimately known to us. We cannot doubt that we are conscious and have a conscious experience like a pain. Yet what is meant by 'knowing' here? What is meant by saying that we *know* how things are with us? When I say that I know that there is a red rose in the garden, then I say so because I have seen one. But when I say that have a pain, I have not seen anything. Conscious experiences like a pain are not objects populating the mental, inner world, for experiences are not entities of any kind. And while we have eyes with which we can observe objects in the outer world and fingers with which we can point to them, there are no 'inner eyes and fingers' with which we can introspect and point. This has a far-reaching consequence: if there are no such things as mental objects to which we can refer, the distinction between appearance and reality does not apply here either. This distinction only applies if we can distinguish recognizing something correctly from recognizing it incorrectly, correctly from incorrectly identifying something, can improve conditions of observation enabling us to see or hear things better ('More light!' or 'Speak louder please'), and so on (see Baker and Hacker, 2005, chapter 5; Smit, 2015). We can make mistakes and correct them when we communicate about colours visible in the outside world, but we *cannot err* with regard to the inner world. One cannot say 'I know I have a pain because I feel it', because feel does not signify a form of perception. It makes no sense to say 'I thought that I felt a pain, but I was wrong' or 'I had a pain but I did not feel it'. Hence the distinction between doubt and certainty does not apply to the mental realm either. There may be indeterminacy (is it just an ache or a pain?) or indecision (I may be unsure about what I think of something), but there is no room for ignorance, for it only makes sense to talk of ignorance if it also makes sense to talk of finding out, coming to know, or learning. But we are not ignorant of pain, do not find out whether we have a pain, and do not come to know this. And while we can answer the question 'How do you know?' when we discuss issues in the outside world by adducing evidence or a source of knowledge, such as perception, there is no such thing as adducing evidence as a response to 'How do you know that you are in pain?'. When I say that I have a pain, adding 'I know' does not add anything in most circumstances. I add this phrase ('Yes, I know I am in pain!!!') as an empathic (not epistemic) remark when someone irritates me because he or she keeps on saying to me that I have a pain. So 'I know I have a pain because I feel it' amounts to 'I know I have a pain because I have it', i.e. because I am not lying (it is true means here that I am sincere). And that is nothing more than saying that 'I have a pain'.

Thus, in contrast to what Tononi and Koch believe, our expressions, avowals of mental phenomena are *non-cognitive* utterances. In the first-person case ('I have a pain') doubt is not excluded because the *epistemic* conditions of certainty are met (as Tononi and Koch think); doubt *and* certainty are here precluded

for logical or conceptual reasons: our verbal utterances, avowals and reports of immediate experiences *have no grounds at all* (Hacker, 2019; Wittgenstein, 2009). We can express our experiences, but we must not (as Tononi and Koch do) confuse the capacity *to say* with the capacity *to introspect or apperceive*. Instead of asking the Cartesian epistemic question of how first-person propositions are grounded in private experiences (there is no such thing), we should pose a different question: what are the conceptual preconditions for a human being to avow or report *groundlessly* how things are with him or her.

Thus there is not any reason to construct explanations by starting from conscious experiences which we know with absolute certainty. We add here that we do not deny that there is a difference between feeling a pain and the physiological changes occurring when a creature has a pain without having these subjective experiences. What we deny is that there is an inner world inhabited with experiences, feelings, sense-impressions, etc. which we name by private, mental definitions, observe *in foro interno*, and describe in words for the benefit of others (we assume that language is a vehicle of our private experiences and thoughts). The problem with this picture, we have already seen, is that there is no inner, mental world populated by mental objects, events, states or processes which we can introspect or apperceive. This has an important consequence: if there is no such thing as a separate, mental realm, then 'the world of consciousness' is, in contrast to what Tononi and Koch believe, not a private world at all. The 'content' of this world can neither be seen by others, nor be seen by me. We have no grounds or evidence for claiming to be conscious. It is not intelligible to say 'I am certain that I am conscious', for 'it seems to him that he is conscious but he is mistaken' is nonsense. For sure, we sometimes say 'I am conscious', but this is never meant as an epistemic claim. We utter this sentence when we for example recover from an anaesthetic, and by this linguistic expression signal to the nurse our recovery. We might just as well have said 'Hello' (Wittgenstein, 2009, par. 416).

Conclusion

We have discussed the differences between the (neo) Cartesian and Aristotelian conception of consciousness and if and how we can integrate them with evolutionary theory. Our investigations have revealed that the (neo) Cartesian conception is incoherent: Descartes identified thinking with consciousness or conscious experiences and argued that consciousness belonged to a separate realm. We have argued that the Cartesian conception leads to unresolvable problems because it is incoherent to argue that there is an inner, mental world populated by mental objects, events, states or processes which we can introspect or apperceive. Hence it is time for a change in the *explanandum*. What we have to investigate is not how expressions like 'I think' or 'I am conscious' are rooted in introspection or apperception, but how on top of the vegetative and sensitive *psuchē* in humans alone the rational *psuchē* evolved during evolution (and evolves during development). That requires that we have to study how the ability to use these expressions emerged (and nowadays emerge when children learn to participate in the normative practice of using words and later sentences), for doing things with words leads to an expansion of our mental powers.

We have discussed why the neo-Aristotelian conception extended with evolutionary theory has far better prospects. This raises a question. Why was and is it hard to see this alternative? One complicating factor is that Descartes's ideas are rooted in breakthroughs in physics (recall that physicists discovered that we can explain matter in motion in the inanimate world without having recourse to Aristotle's concept of purpose). Descartes argued that animals, with the exception of humans, are complex machines understandable in terms of the laws of matter in motion. Aristotle's concept of the *psuchē* is a biological concept. He used the concept of a *psuchē* to demarcate the animate from the inanimate. The *psuchē* is the set of powers (or potentialities) the exercise of which is characteristic of the organism. We have argued that, for understanding the relevance of the Aristotle's conception, we should reconcile Aristotle's concept of the mind with the thermodynamic notion of life. Organisms evolved out of open thermodynamic systems that have acquired heredity. We have elaborated how we can explain how—what Aristotle called—the vegetative, sensitive and rational *psuchē* evolved after the transition from cells to multicellular organisms. We have also explained that perceptual consciousness evolved in animals when they were capable of displaying more and more complex forms of sensitivity to the environment, more complex forms of learning, and more complex forms of goal-directed behaviour. One context in which perceptual consciousness may have evolved is predator–prey interactions. Self-consciousness in humans evolved when they acquired, among other things, the sophisticated linguistic power to apply psychological predicates to themselves without criteria and to others on the basis of behavioural criteria. We expect that the neo-Aristotelian conception extended with evolutionary theory is capable of testing future hypotheses generated within the confines of this framework.

Data availability

Data sharing is not applicable to this paper.

Received: 6 April 2020; Accepted: 3 September 2020;
Published online: 15 September 2020

Notes

1 What we call the neo-Aristotelian conception is not a form of functionalism. Functionalism explains psychological attributes in terms of inputs (stimuli) and outputs (behaviour) mediated by functional operations in the brain. Believing, wanting, perceiving, and so on, are explained as functional states mediating stimuli and behavioural outputs. Neo-Aristotelians deny that psychological attributes are defined by inputs and outputs alone. They explain psychological attributes in terms of the capacities of living, sentient beings expressed in their behaviour.

2 A photon of an appropriate wavelength causes a conformational change in rhodopsin and photopsin, triggering a biochemical cascade that culminates in hyperpolarization of the plasma membrane and discharge of (rod and cone) cells. The point to notice is that rhodopsin and photopsin are produced as the result of transcription and translation of a gene, but the gene does not determine the multiple states of the protein. These states solely depend on the response to photons (environment). Hence the protein is capable of adopting different states, a capacity that the gene itself lacks. One can argue that the overall response of the organism (initiated by the pattern of discharge of retinal cells and resulting in a perception and motor response of the whole organism) is in this sense independent of the genes (see also Haig, 2008). Since this capacity was already present in single-cell organisms during the early stages of evolution, it is the beginning of the evolution of different forms of sensitivity to changes in the environment.

3 There are many examples illustrating how out of an extension of the number and states of proteins, of cells and their connections, et cetera, the sensitivity to the environment expanded (see Sanes and Yamagata, 2009). A well-studied example illustrating how mice become sensitive to their environment is the development of the organization of the cortex involved in navigation (Catania, 2002; see also Kirschner and Gerhart, 2005). Mice use whiskers, a sense organ for navigation, but these do not connect directly to the cortex. There are several structures in between, such as the thalamus, but the point to notice that each whisker has, mediated by interneurons, a one-to-one correspondence to the nerves in the cortex. During the early stages of development there is a critical period in which the innervation of the cortical nerves by signals coming from the whiskers is essential. Experiments have shown that damage during the critical period of one of the whiskers causes elimination of the cortical nerves corresponding to the whisker. Their space is then taken up by cortical nerves corresponding to another whisker. Hence after the cortex, during the early stages of ontogenesis, first develops as the result of local interactions between cells, the axons of

the thalamus (that are indirectly activated by the whiskers) activate during the critical period the cortical region, resulting in a certain cortical organization. However in some species (for example hamsters) the plasticity is maintained for a longer period, explaining why cortical areas can partially be reorganized late in life, just as in humans some patients can recover from brain injury and stroke. Cortical plasticity also explains why in congenitally blind humans some parts of the visual cortex become responsive to tactile input helping the blind to read braille.

4 Cheetahs are an example: they try to capture prey by surprise and are rarely successful against preys that have been alerted to their presence. By contrast: wild dogs hunt like coursers, trotting towards their prey in full view. Interestingly, gazelles stot (leaping off the ground with all four legs held stiff and straight) when they see wild dogs, but rarely in response to cheetahs (FitzGibbon and Fanshawe, 1988). Since wild dogs run down their prey in long chases, it is thought that stotting signals to the hunting dogs that they are healthy enough to outrun them.

References

Baker GP, Hacker PMS (2005) Wittgenstein: meaning and understanding, 2nd edn. Basil Blackwell, Oxford, revised by Hacker PMS; 1st edn 1983

Bennett MR, Hacker PMS (2003) Philosophical foundations of neuroscience. Blackwell Publishing, Oxford

Ben-Yami H (2015) Descartes' philosophical revolution: a reassessment. Palgrave Macmillan, Basingstoke

Catania KC (2002) Barrels, stripes, and fingerprints in the brain—implications for theories of cortical organization. J Neurocytol 31:347–358

Chalmers D (1996) The conscious mind. Oxford University Press, Oxford

Crick F (1995) The astonishing hypothesis; the scientific search for the soul. Touchstone, New York

Dehaene S (2014) Consciousness and the brain; deciphering how the brain codes our thoughts. Penguin Books, New York

Deheaene S, Lau H, Kouider S (2017) What is consciousness and could machines have it? Science 358:486–492

Dennett D (1993) Consciousness explained. Penguin, Harmondsworth

FitzGibbon CD, Fanshawe JH (1988) Stotting in Thomson's gazelles: an honest signal of condition. Behav Ecol Sociobiol 23:69–74

Gazzaniga MS (2018) The consciousness instinct: unraveling the mystery of how the brain makes the mind. Farrar, Straus and Giroux, New York

Gehring W (2014) The evolution of vision. WIREs Dev Biol 3:1–40

Ginsburg S, Jablonka E (2019) The evolution of the sensitive soul; learning and the origins of consciousness. MIT Press, Cambridge

Hacker PMS (2012) That sad and sorry history of consciousness: being, among other things, a challenge to the 'consciousness-studies community'. R Inst Philos Suppl 70:149–168

Hacker PMS (2013) The intellectual powers: a study of human nature. Wiley-Blackwell, Chichester

Hacker PMS (2019) Wittgenstein: meaning and mind, 2nd, revised edn, first edition 1990. Wiley-Blackwell, Chichester

Haig D (2008) Conflicting messages: genomic imprinting and internal communication. In: d'Ettorre P, Hughes DP eds. Sociobiology of communication: an interdisciplinary perspective. Oxford University Press, Oxford, pp. 209–223

Hamilton WD (1964) The genetical evolution of social behaviour; I & II. J Theor Biol 7:1–52

Kirschner M, Gerhart JC (2005) The plausibility of life; resolving Darwin's dilemma. Yale University Press, New Haven

Maynard Smith J, Szathmáry E (1995) The major transitions in evolution. W.H. Freeman, New York

Michod RE (2007) Evolution of individuality during the transition from unicellular to multicellular life. Proc Natl Acad Sci USA 104(suppl 1):8613–8618

Nagel T (1974) What is it like to be a bat? Philos Rev 83:435–450

Rundle B (1997) Mind in action. Clarendon Press, Oxford

Sanes JR, Yamagata M (2009) Many paths to synaptic specificity. Annu Rev Cell Dev Biol 25:161–195

Shannon CE, Weaver W (1963) The mathematical theory of communication. University of Illinois Press, Urbana

Smit H (2014) The social evolution of human nature: from biology to language. Cambridge University Press, Cambridge

Smit H (2015) Popper and Wittgenstein on the metaphysics of experience. J Gen Philos Sci 46:319–336

Smit H (2016) The transition from animal to linguistic communication. Biol Theory 11:158–172

Smit H (2018) Inclusive fitness theory and the evolution of mind and language. Erkenntnis 83:287–314

Smit H (2020) The Cartesian conception of the development of the mind and its neo-Aristotelian alternative. Biol Theory 15:107–120

Smit H, Hacker PMS (2014) Seven misconceptions about the mereological fallacy: a compilation for the perplexed. Erkenntnis 79:1077–1097

Tononi G (2005) Consciousness, information integration, and the brain. Prog Brain Res 150:Ch 9

Tononi G, Koch C (2015) Consciousness: here, there and everywhere? Philos Trans R Soc B 370:20140167. https://doi.org/10.1098/rstb.2014.0167

Tooby J, Cosmides L (1992) The psychological foundations of culture. In: Barkow JH, Cosmides L, Tooby J eds. The adapted mind. Oxford University Press, Oxford, pp. 19–36

Weaver W (1949) The mathematics of communication. Sci Am 181:11–15

White AR (1964) Attention. Blackwell, Oxford

Wittgenstein L (2009 [1953]) Philosophical investigations (trans: Anscombe GEM, Hacker PMS, Schulte J), revised 4th edn by Hacker PMS, Schulte J. Wiley-Blackwell, Chichester

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.S.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.