Scientia*et*Fides 5(2)/2017 ISSN 2300-7648 (print) / ISSN 2353-5636 (online)

Received: November 24, 2016. Accepted: April 22, 2017 DOI: http://dx.doi.org/10.12775/SetF.2017.023

Soul or Mind? Some Remarks on Explanation in Cognitive Science

JÓZEF BREMER SJ
Jesuit University Ignatianum in Krakow
zjbremer@cyf-kr.edu.pl

> But first, a word! Till now I've had no direction,
> When old or young teased me with a question.
> For example: no one's found out, ever,
> What makes body and soul fit together:
> Stick tight, as if there'll be no separation,
> Yet always cause each other irritation
> (Goethe 2003, 6892–6896).

Abstract. In the article author analyses the extent to which it is possible to regard the Aristotelian conception of the soul as actually necessary and applicable for modern neuroscience. The framework in which this objective is going to be accomplished is provided by the idea of the coexistence of the "manifest" and "scientific" images of the world and persons, as introduced by Wilfrid Sellars. In subsequent sections, author initially formulates an answer to the questions of what it is that Aristotle sought to explain with his conception of soul as formal cause, and how this notion could be portrayed in terms drawn from contemporary neurological science. It is author's intention to show that no other concepts—be they the Cartesian "mind" or a Chalmers-style "self"—come anywhere close to matching the breadth of scope of Aristotelian "soul". At the same time, though, analysis carried out here do not intend to undertake any kind of defence of the soul as a spiritual entity of the sort popular where religious faith is concerned,

namely as an entity separable from the body, immortal, and inimitably tied to God. Then, in the second major part of the article, author analyses the extent to which the Aristotelian conception of the soul can be thought of as involving criteria of the sort that we encounter in the contemporary theory of strong emergence. According to the author the theory of strong emergence has, as a heuristic theory, quite general predictive and explanatory power as regards the field of human thought and behaviour.

Key words: soul, mind, self, brain, neuroscience, strong emergence, Aristotle

Introduction

Modern philosophy and science, especially in their more recent and contemporary forms, have proved somewhat ungracious where the soul is concerned. On the one hand, in the Anglo-American philosophical tradition we may observe an increasing tendency—beginning with Locke, Hobbes and Hume—to subordinate the mental (the mind) to the physical (the brain). On the other, we may also note that the Cartesian dualist conception of minds and mechanical bodies brought with it a reduction of the Aristotelian biological conception of soul (as a principle of living beings constitutive of the very form of the body) to the idea of the thinking soul. According to this reduction, our understanding of the living organism is to be modelled on that of dead bodies. The response to Descartes' ontological construal of the *res cogitans* was such as to engender the body-mind problem—one that remains unresolved today; however, it also lent support to dualism, which has come to be regarded as a theory that is scientifically hopeless: the study of life pursued in the expanding fields of biological and neuroscientific research is not permitted to involve any metaphysical objects unverifiable through empirical research or analysis. The vehemence of the dispute about vitalism, which—in terms similar to those of the mechanics of fundamental forces—postulates an independent species of living power, bears witness to this Cartesian mode of thinking.

At the same time, we are entitled to ask the following question: had the soul been understood in accordance with Aristotelian rather than the Cartesian notions, would contemporary philosophy and science now be more accommodating in their attitudes towards it? Taking the example of biology,

we may then assert that as a science it should first constitute itself in a quite new way, as something distinct from physics—by making clear that living bodies cannot be understood solely on the basis of the model furnished by dead corpses. Indeed, discussions concerning the epistemological and ontological characteristics of living versus dead bodies continue to be pursued with vigour, especially in relation to the position that holds that even in the absence of any living power, the biological regularities involved cannot be wholly reduced to physical ones (cf. Krohs and Töpfer 2005).

It seems that commentators do occasionally feel compelled either to couch our current philosophical or scientific explanations in terms drawn from Aristotle's ideas about explanation, or to read Aristotle as anticipating present theories in these or other fields. We hear, for instance, such expressions as "an Aristotle for the age of neuroscience", and, at the same time, encounter talk of the soul conducted in pretty similar terms—seeking, in effect, to secure a place for Aristotelian concepts in our contemporary understanding on the basis of how they show up in contexts furnished by modern scientific approaches themselves (cf. Goodman and Caramenico 2010).

The goal of the present article is to analyse the extent to which it is possible to regard the Aristotelian conception of the soul as actually necessary for modern neuroscience. The framework in which I aim to accomplish this will be provided by the idea of the coexistence of the "manifest" and "scientific" images of the world and persons, as introduced by Wilfrid Sellars. In subsequent sections, I shall initially formulate an answer to the questions of what it is that Aristotle sought to explain with his conception of soul as formal cause, and how this notion could be portrayed in terms drawn from contemporary neurological science. I expect to show that no other concepts—be they the Cartesian "mind" or a Chalmers-style "self"—come anywhere close to matching the breadth of scope of Aristotelian "soul". At the same time, though, I do not intend to undertake any kind of defence of the soul as a spiritual entity of the sort popular where religious faith is concerned—as separable from the body, immortal, and inimitably tied to God. Then, in the second major part of the article, I shall analyse the extent to

which the Aristotelian conception of the soul can be thought of as involving criteria of the sort that we encounter in the contemporary theory of strong emergence (construing the latter as essentially a heuristically validated theory, whose purpose is to stimulate further investigation for the sake of empirically based theorizing.)

1. Aristotle as a philosopher of the manifest image of the world

In his article "Philosophy and the Scientific Image of Man", Sellars introduced his telling description of the "manifest" and "scientific" images, together with some compelling ideas about the relationship between them and what expectations we might have for their reconciliation. Along with this account, he introduced a particular kind of opposition between two conceptions of ourselves and our place in the world.

Taking Sellars' concept of the manifest image as our first point of departure permits us to initially state that a human person is not to be thought of as a Cartesian mind, but rather as a living organism understood in Aristotelian hylomorphic terms. Sellars himself says that "[t]he Aristotelian-Strawsonian reconstruction is along sounder lines" (Sellars 1974, 240), and indeed, it does seem better suited to capturing the crucial structural features of our manifest-image conception of persons. According to him, Strawson achieved a kind of assimilation of the Aristotelian reality of persons within the terms of our contemporary philosophical discourse (cf.: Sellars 1992, 170–171; Rosenberg and Sellars 2007, 13; Bremer 2007, 9–20). Aristotelian persons, as described in Aristotle's "On the Soul", and Strawsonian persons as outlined in "Individuals", are more or less homogenous.[1] They consist of multi-layered properties, but they do not correspond to the dualists' idea of a body-mind tandem as exemplified in

[1] Therefore, what is primitive (original) is not the concept of a mind, but that of a person: "What I mean by the concept of a person is the concept of a type of entity such that 'both' predicates ascribing states of consciousness 'and' predicates ascribing corporeal characteristics, a physical situation, etc., are equally applicable to a single individual of that single type" (Strawson 1959, 101–102).

the Cartesian conception of the person. It is the same person who thinks and who moves. According to the manifest image, persons are fundamental, single individuals. They were a universal category of the Sellarsian (primitive) "original image", about which "[t]he truth is, rather, that *originally* to be a tree was *a way of being a person*" (Sellars 1963, 10, cf. 7–8). When, in the original manifest image, man ceased to think of what we call trees as persons, this change was more radical than a change in belief: it was a change in categorical understanding.

It is the manifest image that has determined "the framework in terms of which man came to be aware of himself as man-in-the-world" (Ibid., 40). In this sense, the primary objects of the manifest image are persons, who reflectively understand themselves as being in the world both as thinkers and as agents. In terms of the manifest image, persons are capable of feeling and perceiving, they are cognitive knowers of the world, they are agents able to affect their—social and natural—surroundings through deliberate and rational forms of elective behaviour. Explanations, we might add, when given in the context of the manifest image, usually follow the model of Mill's inductive approach.

The "scientific image", by comparison, is a complex projection of man-in-the-world and human understanding. It emerges with the ongoing findings resulting from theoretical explanation and reasoning. In particular, it benefits from those approaches that involve postulatory forms of theory construction. The scientific image is, admittedly, methodologically dependent on the categories of the manifest world, but

> it purports to be a *complete* image, i.e., to define a framework which could be the *whole truth* about that which belongs to the image. Thus although methodologically a development *within* the manifest image, the scientific image presents itself as a *rival* image. From its point of view the manifest image on which it rests is an "inadequate" but pragmatically useful likeness of a reality which first finds its adequate (in principle) likeness in the scientific image (Sellars 1963, 20).[2]

[2] In what I consider to have been a highly inspiring discussion that took place during the International Workshop "Soul or Brain: What Makes Us Human" (Toruń, 2016), the question arose of whether the solution to the body-mind problem proposed by Peter Hacker

The scientific image of man-in-the-world is multidisciplinary. It is partially elaborated by neuroscience. A person is seen as composed from sub-personal parts, consisting of a great many billions of neurons connected by many trillions of synapses—a brain that itself has no indication of what is going on amongst those synapses. There are several possibilities for how we should understand the connection between the manifest-image and scientific-image concepts of a person. In manifest-image terms, I or he/she is (a) either ontologically or epistemologically reducible to the brain, or (b) strictly correlated with the brain, or (c) separated from it. But which of these ways of understanding a person, if any, entails an adequate reconciling of the subpersonal entities involved with the manifest image of the conscious me as me—caught up in my particular life, recalling it, planning for it? How can the view from the neurological stage be reconciled with the one that shows

---

could be adapted, together with Bennett and Hacker's critique of the so called "mereological fallacy" (i.e., the inclination to attribute to the brain mental concepts that only make sense when attributed to the whole person). (See: Hacker and Bennett 2003). Bennett is a neuroscientist, while Hacker is a philosopher, known primarily for his commentary on the later philosophy of Wittgenstein. In Hacker's view, philosophy is an *a priori* discipline, in that its assertions or statements are conceptually true (Hacker and Bennett 2003, 3, 6, 318). But what does "conceptual true" actually mean? In the context of a conjunction of philosophy with science, the distinction between *a priori* and *a posteriori* surely needs to be supplemented by the analytic-synthetic distinction. A synthetic statement is one whose truth is determined partially by the meanings of the terms used, and partially by the reality of how things stand in the world. The abovementioned book, moreover, presents a very limited perspective on philosophy. A more revealing title for it would have been "A Wittgensteinian Approach to the Conceptual Elements of Cognitive Neuroscience". Philosophy, to my mind (and also in line with Sellars), is to be more broadly conceived than just what we find in late-Wittgensteinian philosophical approaches to (ordinary) language construed in manifest-image terms. Hacker and Bennett argue that just about everybody in cognitive neuroscience is guilty of committing a rather simple conceptual "mereological fallacy". But are they offering any alternative that might be of benefit to the neurosciences? Are they presenting any positive theory or model, such as would shed light on how such neuroscientific theories or models might be structured? For example, their reports about commissurotomy (Sperry, Gazzaniga), correct as they may be, essentially just consist of an accumulation of redescriptions of accessible, manifest phenomena: they are not explanations at all. Hacker and Bennett are surely right: these phenomena occur under conditions that are far from ordinary. But how are we to explain them? Is it sufficient—within the terms of the scientific image as we understand it today—to declare, along with Wittgenstein, that "[e]xplanations come to an end somewhere"? (Wittgenstein 1953, §1).

up as true at the level of the manifest image, and which says that at some point our brains must be correlated with our own thinking about ourselves and about things? Could the latter's aboutness be simply determined by what goes on in and amongst that huge cluster of neurons and synapses?

Sellars requires that we seek an answer to several questions. Is it possible to reconcile these two images of a person? Could manifest-image persons be reduced to such systems of scientifically postulated objects as groups of neurons? Are manifest-image persons ultimately real, and scientific entities merely abstract, in the sense of being posited constructs valued only for their role in the prediction, explanation and control of actually perceived manifest-image persons and objects? Or are manifest-image objects appearances to human minds of a reality constituted by systems of imperceptible particles? Sellars himself opts for a Kantian approach: the manifest image is the phenomenal realm, whereas science reveals things as they are in themselves. On the one hand, admittedly, he speaks of "the primacy of the scientific image" (Sellars 1963, 32), but on the other, he embraces in the end the argument for a "synoptic vision" (Ibid., 40), according to which the descriptive and explanatory capacities of the scientific image are united with the social language of the community and with our individual intentions. That language "provide[s] the ambience of principles and standards (above all, those which make meaningful discourse and rationality itself possible) within which we live our own individual lives" (Ibid.). And this discursive framework, according to Sellars, does not need to be reconciled with the scientific image, but rather could just be "joined to it" (Ibid.; cf. Hodgson 2012, 47, 54).

In the context of some more detailed considerations, Sellars considers and rejects Descartes' attempt to integrate the manifest and scientific images (cf. Sellars 1963, 25–31). On the one hand, on Sellars' reading, Descartes denied that there were brain states which were, in the same sense, the cerebral counterparts of conceptual thinking, while on the other he stated that we know what thinking is without conceiving of it as a complex neurophysiological process, from which it follows that it cannot itself be a complex physiological process. As Sellars himself notes,

at the time of Descartes, theoretical science had obviously not arrived yet at the neurophysiological level of inquiry—except in the form of a clumsy promissory note.

2. Aristotle on types of causes

Sellars recognises the extent to which Aristotle's hylomorphic metaphysics represents a forceful interpretation of the fundamental logic and science of the manifest image: "It should be clear that I regard Aristotle as the philosopher of the Manifest Image" (Sellars 1975, 303). According to Aristotle, "science" means "causal knowledge": in other words, knowledge about what causes are is crucial for every form of science. We only have knowledge of something when we have understood its causes.

> Since we think that we understand something when we know its explanation, and there are four sorts of explanation (one, what it is to be something; one, that if certain items hold it is necessary for this to hold; another, what initiated the change; and fourth, the purpose), all of them are proved through the middle term (Aristotle 1993, 94a20).

This might be summarized as follows: in the framework of the manifest image we demand to know the essence (the "suchness"), the necessary conditions, the (efficient) cause, and the purpose. In this sense, Aristotle's "causes" are often better thought of as "explanations" or "reasons". In his *Metaphysics*, Aristotle refers us to an artist's or a craftsman's work to describe the four types of causation that form the basis of the natural world of the manifest image.

The material cause is "that out of which" the object (statue) is made, the efficient cause is the source of the object's principle of change or stability, the formal cause is the essence of the object, and the final cause is the goal/end of the object, or what the object is good for (cf. Bremer 2015, 101–102). Matter of some kind or other (e.g., gold) has the potential (based on its "plasticity") to be actualized (thus "formed", "organized") by form in

different ways. Aristotle also saw a correlation between formal causation and personal causation (on the part of an agent or soul):

> The mover or agent will always be the vehicle of a form, either a 'this' or 'such', which, when it acts, will be the source and cause of the change, e.g. the full-formed man begets man from what is potentially man (Aristotle 2009, 202a9–11).

Likewise, Phidias the sculptor is the agent cause who arranges some part of the available gold for the purpose of creating the colossal sacred statue of Athena Parthenos. This means that agent causes can be construed in three different senses: as efficient, formal or final causes, which often coincide when we think about the nature of agent causation. The final cause is not principally a causative, but rather an influential cause, because it is only through its having been imagined by Phidias that it has played its role in the genesis of the statue. However, we can still reduce this mental goal to a causative one, if we reflect on the counterfactual consideration that had the artist not thought of and intended some image, he probably would not have become involved in executing the various movements that he did in fact perform. Moreover, if his movements had been wholly arbitrary, they definitely would not have resulted in such a perfect resemblance of Athena. Thus we have here a kind of partial causation. In the example we have given, the formal and final causes are identical: a certain shape resembling that of Athena. At the same time, Sellars reminds us that

> neither Aristotle nor the Scholastics conceived of soul and body as two changing, interacting substances, as did Plato and Descartes. The Aristotelians define the human soul as the substantial form of the human individual, who consists of matter and form as does any terrestrial substance. With qualifications which will be introduced at the proper place, it is the human individual that acts, that does things, and not the soul; just as it is water that freezes, and not the nature of water (Sellars 1949, 558–559).

Souls really exist, but in this broadly outlined Aristotelian sense, not a Platonic or Cartesian one. They possess reality at a different level from that

of any material or spiritual stuff. The reality of souls is properly captured by the concepts of form and formal causation, and these concepts need to be accommodated—in some reasonable way—within the scientific image of persons.

The individual agent mentioned above issues from a hierarchically organized biological system, exhibits causal powers, and counts as a fundamental entity in its own right. "What is ensouled is distinguished from what is not ensouled by living" (Aristotle 1986, 413a22).

The formal causation that figures in the soul's relation to the living body is a species of something metaphysically fundamental, rather than just some sort of efficient causation. In this sense, the soul acts on the body merely in a mediated way, using bodily parts. No direct causation from soul to body is required, and this in turn means that there also need not to be either any new fundamental force involved, or any threat to the preservation of causal laws. The nomological integrity of the microphysical (or neurological) level in no way jeopardizes the real and irreducible power of the whole organism, as Aristotle's hylomorphic approach involves no assumption to the effect that the microphysical facts are somehow elementary and unfounded. It is the form (soul) that is (partially) accountable for the relative places and trajectories of the microphysical parts. Any laws instantiated at a microphysical level assume, as inputs, something already existing within the context of the nature of the whole living, rational, individual human person.

3. Aristotle's soul and strong emergence

The Aristotelian concept of "form" (i.e., *morphê*, meaning a cause which is not itself a material part of the body, but is active in its development and determines its principal features) is in fact directly related to the modern, scientific conception of "information" as something transmitted across generations (cf. Deacon and Koutroufinis 2014, 404). The living body is a specific kind of "in-formed" matter. Even the teleological perspective of Aristotelian physics, which seemed to have been decisively eliminated by the mechanistic paradigm of Cartesian and Newtonian approaches, has

been partly rehabilitated by present-day biologists and economists, and has come to be employed as a basis for a new ecological ethics and economics.[3]

Does any information present in matter serve to bring about the emergence of the individual agent or subject? Aristotle would say that information in a biological system should be regarded as furnishing an additional constraint on a theory of the emergent acting subject. Such a subject is "produced by" (generated by) a hierarchically organized biological system, exercises causal powers, and counts as a fundamental entity in its own right. "What is ensouled is distinguished from what is not ensouled by living" (Aristotle 1986, 413a22). And the "whole" is "something over and above its parts, and not just the sum of them all" (Aristotle 2017, 1045: 8–10). A decisive consequence of this point is that the effects produced by "wholes" offer an unambiguous answer to the charge that wholes are solely "epiphenomena": i.e., that they are nothing more than an expression of their parts, without any causal efficacy. Briefly, a whole exists when it acts like a whole, when it produces (generates) combined effects that the parts cannot produce (generate) by themselves.

Souls are introduced by Aristotle as a kind of explanatory posit in the context of the manifest image of the world. First of all they explain the fact that some beings are alive (vegetative, sensitive, and rational) and others are not. At the same time, he puts the explanatory focus on the type(s) of soul that an organism possesses, insisting that plants have a vegetative soul (explaining reproduction and growth), animals a vegetative and sensitive one (explaining mobility and sensation), and humans one with three powers, in that it is vegetative, sensitive and rational (able to think and reflect).

Thus, all of the activities of the organism are to be explained by means of the functions of its soul.

> The soul is the cause or source of the living body. The terms cause and source have many senses. But the soul is the cause of its body alike in all three senses

[3] The Indian economist and philosopher, Amartya Sen, for example, invokes Aristotle as one of the sources for his ideas (Sen 2000, 289).

> which we explicitly recognize. It is (a) the source or origin of movement, it is (b) the end, it is (c) the essence of the whole living body (Aristotle 1986, 415b10).

Soul is real (i.e., it can act), though not as much so as an independent, separable entity. The representative active constituent of reality—that which is disposed to exhibit causal efficacy in respect of the surrounding world—is the soul-body whole that is an individual person (the man). Aristotle rejects the idea that the soul is the subject of psychic properties:

> Yet to say that it is the soul which is angry is as inexact as it would be to say that it is the soul that weaves webs or builds houses. It is doubtless better to avoid saying that the soul pities or learns or thinks and rather to say that it is the man who does this with his soul (Witt 1992, 180; cf. Aristotle 1986, 408b12–13).

Aristotle avoids here the mereological fallacy of ascribing to parts attributes that can only intelligibly be ascribed to the wholes to which those parts belong. He argues for a sort of holism rather than dualism, as something entailed by the concept of the soul as the principle of realisation of the living body. Soul is not only the organized and functioning whole, but also its directing part.

Ipso facto—unlike in the case of Descartes—conceived in Aristotelian terms, animals' bodies, and human bodies inasmuch as they are similar to them, cannot in principle be explained in purely mechanical terms. Descartes supposed that the nervous system functions like a complex hydraulic machine. Because the structure of our body and the patterns of our behaviour are similar to the structure and behaviour of some animals, it seemed obvious that many of our actions could also be amenable to an underlying mechanistic explanation (cf. Bremer 2014, 7–26).

Pursuing Descartes' approach to the scientific image of the person, along with the various explanations of the mind-body problem proposed in the context of this, has the consequence that we find ourselves confronted with strictly reductionist theories, along with dualists and non-reductionists (cf. Bremer 2010). We may nevertheless wonder whether there is a sense in which Aristotle's manifest-image account of explanation could be shown

to link up with contemporary approaches to explanation of the sort encountered in attempts to determine the scientific image of the world and of persons—e.g., where the latter's thoughts and actions are concerned? Could the abovementioned internal features of Aristotle's theory of soul, and its understanding of it as emergent in the sense of being "produced by the whole", just amount to some sort of contemporary notion of weak emergence? According to David Chalmers:

> A high-level phenomenon is *weakly emergent* with respect to a low-level domain when the high-level phenomenon arises from the low-level domain, but truths concerning that phenomenon are unexpected given the principles governing the low-level domain. Weak emergence is the notion of emergence that is most common in recent scientific discussions. A high-level phenomenon is *strongly emergent* with respect to a low-level domain when the high-level phenomenon arises from the low-level domain, but truths concerning that phenomenon are not deducible even in principle from truths in the low-level domain … I think there is exactly one clear case of a strongly emergent phenomenon, and that is the phenomenon of consciousness (Chalmers 2006, 244).

The concept of weak emergence used in the case of soul-body causation implies that the causal powers of the soul are fully explainable in terms of their being related to the causal powers of the body or its parts (cf. Bremer 2005, 138–139). Strictly speaking, weak emergence relates to unpredictable higher-level states that are ontologically reducible to lower-level states. The ability to reduce higher-level states to their lower-level constituents is useful for setting up a mechanistic explanation of emergent features. Consequently, weak emergence equates with a materialist kind of causal reduction. Such emergence brings entails physical monism: all elements in our manifest world consist solely of physical parts that exist in the scientifically grasped world.

Nevertheless, Aristotle is not in fact a materialist reductionist when it comes to the causal powers of the soul. According to him, an exhaustive description and explanation of the soul-body interaction requires not only material causation, but also the other three types of causation. He

understands soul—in living organisms—as the ontological foundation of efficient, final and formal causation. He realizes that each living thing has just one soul, the actions of which indicate some degree of nutritive, sensitive and (in humans) rational functioning. This soul is the formal, efficient and final cause of the survival of the organism. Only its material cause resides entirely in the body.

Souls are emergent forms of biological organisation, in a sense that denotes "all that makes person as such distinctive—our cognitive and emotive, active and creative, appetitive and moral capabilities" (Goodman and Caramenico 2010, 4).[4] Aristotelian soul is the name for the integrating, organizing capabilities of sentient, sapient and agential personhood.

The emergent person (i.e., the rational, voluntary soul) is an efficient, final and formal cause that is empowered to affect the matter that produces it. *Ipso facto* Aristotle is obliged to assume downward causation, and in consequence of this is committed to a strong kind of emergence. Strong emergence has two characteristics: the novelty of the emergent features entails their irreducibility, and they are able to exert downward ("top-down") causation. Mental states thus have genuinely new causal powers of their own, irreducible to those of the neurological states from which they themselves emerged.

> As we will see, one type of irreducibility seems to imply downward causation while the other seems to imply epiphenomenalism. The failure to keep apart the two kinds of irreducibility has muddled the recent debate about the emergence of properties (Stephan 2002, 86).[5]

Generally, we may say that a systemic property P (e.g., a mental property) of a system S is irreducible, if it does not follow (even in principle) from the behaviour of the system's parts that S has P. Understood this way, the Aristotelian thesis is in line with the contemporary notion of strong emergence, which, according to Stephan, means that

[4] We do not follow the authors in their deliberations about the transcendent dimension of the soul (Goodman and Caramenico 2010, 243).

[5] But see also: Kim 2006, 195–196.

> [w]ithin the physical domain, we would just have to accept additional types of causal influences besides the already known basal types of mutual interactions (Stephan 2002, 89).

The Aristotelian thesis thus requires that we invoke the criteria for strong emergence, and *ipso facto* entails a broader ontology than physical monism: (a) the acting soul is generated by the brain without being reducible to it, because of the unity of its mental experiences, and (b) the acting soul possesses, qua formal cause, the power to affect (i.e., organize) a person's brain, thanks to the latter's neuroplasticity. In the subsequent sections of this article we shall discuss these two specific points.

3.1. The acting individuum: the unity of emergent properties

Along with mental properties located in specialized areas distributed across the brain, a special kind of entity also emerges: the individual and irreducibly singular individuum, equivalent to the Aristotelian agent. In his terms, this is what would be called the "active intellect". Such an individuum possesses its mental properties in a unified manner. The emergent individual explains the unified character of both human experiences and mental properties. On the one hand, an irreducibly singular entity is produced by a sufficiently organized conjunction of areas of the brain, while on the other, that entity unifies the emergent mental properties across diverse brain modalities. According to Aristotle, the acting self (the active intellect) is ontologically distinct from its matter, on the basis both of its irreducibility to that matter and of its unifying nature.

> Intellect in this sense of it is separable, impassable, unmixed, since it is in its essential nature activity (for always the active is superior to the passive factor, the originating force to the matter which it forms) (Aristotle 1986, 430a18–20).[6]

It is the part of the soul responsible for reasoning, known as the active intellect of the soul (*nous poietikos*) and possibly capable of existing separately

[6] Cf.: Bunnin and Yu 2004, 12; Sellars 1949, 561–562.

from the body. In this regard, Aristotle makes a terminological distinction between soul (*psyche*) and intellect (mind, *nous*). Active intellect may survive bodily extinction.

We may add here that philosophers working from the manifest image are not alone in upholding that experience involves an active unifying something (a subject or process). That position is also endorsed by some outstanding neuroscientists engaged in the pursuit of a conception of persons appropriate to the scientific image.

The results of Benjamin Libet's experiments have confirmed such a hypothesis: the readiness potential (RP, evoked unconsciously in the brain) of subjects appears some 500 milliseconds before any reported awareness on the part of a person of a random, spontaneous decision to move their finger. Libet's account makes this quite clear in completely empirical terms, but in philosophical terms the problems were only beginning. Libet's research, it seems, delivers a scientific-empirical argument to the effect that free will is a delusion. But how could we, as unified subjects, reflectively conclude about ourselves that we are responsible for decisions which we were not aware of until after they had been realised?

For sure, some would have been only too happy to see mental states of free will eliminated, but Libet himself was not ready to let them go so unproblematically. Although the unified subject's "decision" to move a finger happened in its brain too early on for it to have been initiated by a conscious, volitional mental state, there was still—just—a timeframe in which conscious awareness might conceivably veto the moving of the finger.[7] This timeframe lasts, on Libet's account, no more than about 100 milliseconds. But is it possible to decide to move at a random moment while at the same time holding on to the view that I will not, ultimately, carry out the movement? The neuroscientist John-Dylan Haynes—together with others—has proposed a kind of experimental proof of the vetoing's occurrence in respect of the acting of a unified subject:

[7] For a detailed description of Libet's experiments, see: Bremer 2013, 194–212.

> To summarize, our results suggest that humans can still cancel or veto a movement even after onset of the RP. This is possible until a point of no return around 200ms before movement onset. However, even after the onset of the movement, it is possible to alter and cancel the movement as it unfolds (Schultze-Kraft et al. 2016, 1084).

Libet himself offers a philosophical solution to both the veto-problem and the issue of the unified something responsible for conscious decisions. However, insofar as his global theory of consciousness is also designed to secure the status of both the acting unified subject and his or her free will, along with the subjective experience which (in his view) carries it, our free decisions must be conscious decisions, and as such these require unified, subjective awareness. He thus postulates the existence of a conscious mental field (CMF), intended to account for the unity of those varied mental phenomena whose unified character would not appear anyway, just as a natural consequence of the firing of neurons.

> the conscious mental field [CMF JB] theory […] is not dualism, in the Cartesian sense; the CMF does not exist without the living brain, and is an emergent property that brain (Libet 2003, 28)

Libet believes that the presence of such a field helps us to explain how the diverse forms of activity of the brain are bound together into a single, unified conscious experience. He does not see this field as a straightforward physical phenomenon, but as an emergent entity. Neither does he assume that manifest conscious volitional activity is explainable by, or reducible to, neuronal activity, although it certainly requires it. Moreover, he does not seek to advocate any kind of dualism or spiritual theory: instead he has proposed CMF as a strongly emergent phenomenon—as something that arises from the combination of active neurons, but which amounts to something definitely greater than and different from the mere sum of all brain activity involved.

The conscious mental field is conceived, then, as an emergent property of brain processing. It exerts an active, unifying and controlling role in

shaping the flow patterns of cerebral excitation. These higher-order mental patterns have their own subjective qualities and operate according to their own causal laws and principles, which are different from and cannot be reduced to those of neurophysiology.

Libet seeks to stave off criticism, noting that some may wish to assign his position to the dualist camp. Yet he is not one to be frightened off by such name-calling. According to him the CMF does not represent the dualism of Descartes, who described the mind as a separable substance. But his CMF proposal is, of course, highly speculative. Meanwhile, he himself claims not to know of any existing evidence that would contradict the proposal, and, furthermore, considers it amenable to direct validation via experimental testing.

Libet points to the results of the neuropsychologist Roger Sperry, who had likewise proposed that the mind is an emergent property of certain appropriate brain functions, equipped with special attributes not evidenced by the neural activities occurring within the system that is the brain itself (cf. Gazzaniga 2013, 2–3).[8] Sperry does clearly hold that mental properties have causal force:

> The conscious subjective properties in our present view are interpreted to have causal potency in regulating the course of brain events; that is, the mental forces or properties exert a regulative control influence in brain physiology (Sperry 1976, 165).[9]

Due to his work with split-brain patients, Sperry initially alighted upon the term "interactionism" to describe his findings. Because these patients' *corpora callosa* had been separated, no neurophysiological account could be given of the unified consciousness they continued to manifest in their ordinary lives. Thus, Sperry argued, there had to be unified interactions at

[8] "In those efforts, I reviewed neuroscientific data which supports the modular view of brain organization, now widely established, along with a possible understanding of why our subjective life seems largely unified. … People —even split-brain patients—feel integrated, whole and purposeful, not modularized and multiple" (Gazzaniga 2013, 4–5).

[9] Cf. Sperry 1987, 164–166.

the emergent level of consciousness. This would then involve conscious states exercising direct downward causation on subsequent brain states (perhaps together with other causal factors). Experiments on split-brain phenomena do not appear to furnish any basis for arguing against either the hypothesis of strong emergence or the notion of a single unified subject with a divided brain.

In this context, M. Gazzaniga, a student and follower of Sperry, highlights a further point:

> On the neurophysiological level, we are born with a sense of fairness and some other moral intuitions. These intuitions contribute to our moral judgments on the behavioral level, and, higher up the chain, our moral judgments contribute to the moral and legal laws we construct for our societies. These moral laws and legal laws on the societal scale provide feedback that constrains behavior. The social pressures on the individual at the behavior level affect his survival and reproduction and thus what underlying brain processes are selected for. Over time, these social pressures begin to shape who we are. Thus, it is easy to see that these moral systems become real and very important to understand (Gazzaniga 2011, 186).[10]

Social pressures (i.e., non-material and non-efficient causes) do *de facto* change behaviour, as is demonstrated by scientific studies carried out in relation to various different versions of the scientific image. Persons make decisions, and they are formed by societies. Thus, claims to the effect that all forms of behaviour are completely controlled by neurological processes in the brain are not supported either by our experiences in respect of the manifest image, or by what scientific research and explanations have shown to be the case.

[10] For a comparison between Sperry's and Gazzaniga's concepts of emergence, see: Seager 2012, 227.

3.2. The unified subject and its formal causal power

> "Todo hombre puede ser, si se lo propone, escultor de su propio cerebro".[11]
> Santiago Ramón y Cajal

The unified subject possesses the causal power to affect (organize) the agent's brain, thanks to the latter's neuroplasticity. Formal causation connected with material, and goal-oriented causation, mean that the emergent conscious agent has the causal power to organize his or her own brain. Some discoveries about the brain's neuroplasticity support the thesis that the brain is capable of acting under the guidance of such a conscious individuum.

> The notion of neuroplasticity, the brain's capacity to change and alter its structure and function, is particularly relevant to rehabilitation and an understanding of both natural and induced recovery processes (Mateer and Bogod 2003, 168).

It is widely acknowledged that after non-progressive brain damage most individuals exhibit some degree of recovery of cognitive and behavioural functions, and that many make a significant improvement. Healthy brain-areas and circuits have the ability to assume new functions or roles—ones usually performed by the damaged parts. The concept of neuroplasticity has come to be used as a broad bridge between the field of neurobiology and different kinds of therapy—above all psychotherapy.

> The study of neuroplasticity is an extremely broad field that is investigated at multiple levels from molecules, to neural systems, to behaviour. Much of the seminal work regarding neuroplasticity has arisen from investigations of the somatosensory and motor systems (Merabet, Lotfi and Pascual-Leone 2010, 44).

In other words, our brains are capable of being formed—according to Gazzaniga—by the absorption of information from the world, but also by conscious subjects themselves. The plasticity of the brain permits conscious

[11] "Every man can, if he so desires, become the sculptor of his own brain" (Llaca-Rodríguez and Ramón y Cajal, 202). Santiago Ramón y Cajal is regarded by some as the father of modern neuroscience.

subjects to actively form neural stimuli, and thus to instigate short-term and long-term changes to the relevant neural circuits fundamental to human actions and to the formation of behaviour, habits and habitual responses (Cheung et al. 2014, 6).

Neuroplasticity, understood in these terms, correlates well with the metaphysical claim underlying Aristotle's hylomorphic approach to the soul-body connection. The neural processes—to which the conscious subject relates—have the capability to be managed (i.e., shaped) by this conscious unified subject, much as a choirmaster manages (i.e., shapes) the musical ensemble of singers into a melodious whole. The subject exerts its (formal) causal influence in order to modify (or rewrite) the neural connections of its own brain.

Such subject-directed neuroplasticity means that the subject has a voluntary capacity to manage (or organize), through sustained conscious effort, the areas of the cerebral cortex, so that new (or even, perhaps, sometimes old) forms of behaviour emerge. Subject-directed neuroplasticity essentially uses:

> the power of focused attention along with the ability to apply commitment, hard work, and dedication, to direct your choices and actions, thereby rewiring your brain to work for you (Schwartz and Gladding 2012, 39).[12]

Brain neuroplasticity, the stimulation of the growth of neuronal connections, and the influence of culture create the social, emergent mind that is a hallmark of human beings. This implies that brain organization is more important than, for instance, brain height or weight, and that the human brain is, *contra* Descartes, organized in a quite different way from the brains of other animals (cf. Gazzaniga 2011, 25).

3.3. Empirical examples of the evidence for subject-directed neuroplasticity

1. According to Research Professor of Psychiatry Jeffrey M. Schwartz, conventional science has long held the position that "the mind" (or the soul) is merely an illusion. According to this view, the mind is an epiphenomenal

[12] Cf. Nahum, Lee and Merzenich 2013, 141–173.

side effect of electrochemical activity within the physical brain (i.e., an effect without any causal powers). Schwartz, though, argues for exactly the opposite thesis: that the mind has a life of its own as an independent entity that can shape and control the functioning of the physical brain (cf. Schwartz and Begley 2002).[13]

Schwartz proposes a non-materialist interpretation of neuroscience, and asserts that such an interpretation is increasingly revealing the limitations of the standard reductionist-materialist interpretation. He underpins this with the results that he himself has obtained as a specialist working in the treatment of—amongst others—patients with obsessive-compulsive disorder. Somebody suffering from this disorder recognizes unwanted obsessive-compulsive thoughts as such, experiencing the urges that stem from them as quite foreign to his or her intrinsic self. Yet such intrusive thoughts (the obsession part) do still trigger intense urges to accomplish ritualistic forms of behaviour (the compulsive part) (cf. Schwartz and Begley 2002, 55–56, 80–87, 355). By way of example, we may mention that after a few episodes handwashing the compulsive person knows perfectly well that his/her hands are clean, and yet still feels compelled to wash them again. The visible difference between the obvious truth (the hands are clean) and certain irrational, unreasonable doubts (to the effect that they might still be dirty) has encouraged Schwartz to examine the philosophical foundations of this area of neuroscience.

Scans of certain brain regions of patients exhibiting obsessive-compulsive disorder have led Schwartz to the thesis that these regions are exhibiting unusual patterns of activity. Taken by itself, such results are consistent with a reductionist-materialist view of mind: after all, materialism implies that it is a consequence of the brain making the mind possible that unusual patterns of brain activity can be presumed to correlate in a necessary and sufficient way with dysfunctional mental states. Nevertheless, having found unusual patterns of brain activity, Schwartz has engaged patients with obsessive-compulsive disorder in intensive, real mental 4Rs-therapy

[13] Because the book mostly describes Schwartz's own research and thinking, I will refer just to him as the primary author of the book.

(relabelling, reattributing, refocusing, and revaluing). In the case of compulsive hand-washing cited above as an example, this therapeutic effort involved a patient recognizing that his hands were indeed clean (relabelling). The urge to wash the hands is seen as not pointing to any reality that must actually be acted on. It is crucial that the patient develops a kind of mindful attentiveness to this truth, and writing down such truths can be extremely helpful, along with attributing fears and doubts about one's hands being dirty to a dysfunctional brain (reattributing), channelling one's reflections and actions away from handwashing and towards profitable goals (refocusing), and, finally, comprehending at a deep level the absurdity of obsessive-compulsive messages (revaluing). The 5thR comes from the side of the brain itself, in the form of a realignment of forms of brain-activity.

Schwartz has found not only that patients who have attended his 4R-s therapy have experienced substantial emancipation from obsessive-compulsive symptoms, but also that their brain neuroimaging scans indicate a sustainable reframing of brain-activity patterns. Thus, without any interventions (such as would directly affect their brains) either from outside (e.g., through transcranial magnetic or transcranial electric stimulation) (Bremer 2013, 246–249) or from inside (e.g., through medical drugs), obsessive-compulsive patients have proved able to organize their brains by deliberately altering their thoughts and behaviours. The significant point for Schwartz here is plainly not that modified thoughts and forms of behaviour were permanently altering the patterns of a person's brain activity, but rather that such modifications were resulting from, as he calls it, "mindful awareness…, that has an activating effect on the brain" (Schwartz and Begley 2002, 335): i.e., conscious and purposive thoughts or actions in which the patient adopts the stance of an uninvolved observer. The therapist's role is to draw the patient's attention to (i.e., focus it on) aspects of his present experience that he might not otherwise notice.

On the basis of such experiments involving therapy, Schwartz argues that since the mind can actually change the brain, the mind must be something other than the brain: something ideal. He shows that the human mind is something autonomous, which can shape and control the functioning of the

physical brain. On the one hand, these experiments could be interpreted as being entirely consistent with mainstream neurology—the material brain is, after all, modifying the material brain. On the other hand, Schwartz's therapy has its basis in our emerging understanding of adult plasticity—it lies in the brain's ability to be radically rewired not just in childhood, but throughout life. Yet in this paradigm-shifting work, the author takes neuroplasticity one critical step forward. Over the course of years spent medicating patients with obsessive-compulsive disorders, Schwartz arrived at an important result: while following the 4R-rherapy that he had developed, his patients were able to remodel their own neural pathways. Audaciously, he suggests that we are more than mere kinetic robots: our physical brain does not shape our destiny alone.

2. The next example of evidence for agent-directed neuroplasticity can be found in cases of strokes.

> Stroke can be seen as a massive distortion of the capacity of the brain to process neural information, with heterogeneous consequences. Not only is the motor system affected after a stroke, but also the cognitive and emotional systems may be seriously impaired (de Vries and Mulder 2007, 5).

According to de Vries, recent papers indicate that information provided by imagining and observing movements might form an additional source of the kind information that could be useful for motor rehabilitation after a stroke. Meanwhile, one of the first conceptualisations of the imagination comes from Aristotle: "The soul never thinks without a mental image [*phantasma*]" (Aristotle 1986, 431a15; cf. Tateo 2015, 1–6). The radical imagination of the singular individual is the essential determining element of the human psyche. It includes images, memories and hopes.

It is fascinating to wonder whether, for stroke victims, information provided by imagining (motor imagery) and observation may also play a role in the process of re-learning motor skills. Motor imagery is the mental execution of a movement without any overt movement, or without any peripheral (muscle) activation. It has been shown that motor imagery

leads to the activation of the same brain areas as the actual movements do. The justification behind this assertion is that brain areas normally involved in movement planning and execution are also active during the imagining of a movement. Most studies of stroke victims combine physical and mental practice, and use auditory instructions to guide patients through the imagined movements. Patients are encouraged to listen to an audiotape instructing them to imagine themselves making functional use of the affected limb.

> Studies measuring electromyographic (EMG) activity, cortical motor evoked potentials, and cerebral blood flow also have shown that the appropriate neuromotor pathways imagined as being used are actually being used and that metabolic activity of neurons is increased during mental practice as if the activity is actually being performed (Page et al. 2001, 1457).

Mental and physical practice at the behest of the individual lead *ipso facto* to neuroplastic changes in the motor cortex area of the brain.

According to Carvalho (et al.), various approaches have been employed to investigate post-stroke rehabilitation (Carvalho et al. 2013). It has been shown that the human brain is capable of significant recovery after such injuries. Among its follow-on effects, hemiparesis (affliction of one half of the body: left or right) has been treated with mirror-therapy, from which cortical changes have been found to ensue. Mirror therapy is a type of motor imagery in which a mirror is used to transfer visual stimuli to the brain through observation of one's uninfluenced body part as it makes a series of movements. The underlying principle is that movement of the affected limb can be stimulated *via* visual cues originating from the opposite healthy side of the body. On this basis, it is thought that such a form of therapy can bring about useful effects in patients who have lost movement of an arm or leg, including those who have had a stroke.

In particular, sensorimotor disorders in post-stroke patients during the execution or observation of motor action have induced changes to the neighbouring cortical semi-shade area. Additionally, motion imagination studies have demonstrated efficacies in treating the post-stroke group of

patients. In this case, the underlying hypothesis is that "mirror neurons" have been activated during such trainings. According to Acharya and Shukla:

> Mirror neurons represent a distinctive class of neurons that discharge both when an individual executes a motor act and when he observes another individual performing the same or a similar motor act (Acharya and Shukla 2012, 118).

Mirror neurons not only fire when we perform a movement, but also when someone else executes that movement and we observe their performance. As such, they are in line with one of the most striking aspects of our complex thought processes: namely, "intention understanding". This reflects the fact that there are two distinct information-based processes one can be on the receiving end of when observing an action done by another individual: the first concerns our coming to know "what" action is being performed, while the second, more complex one concerns our understanding "what for" or "why" (i.e., with what intention) the action is being done. There has also been speculation about mirror neurons possibly providing the neurological basis for human self-awareness. Mirror neurons can not only help simulate other people's behavior, but can also be turned "inward" to create second-order representations or meta-representations of one's own earlier brain processes. Ramachandran, for instance, believes that this could constitute the neural basis for introspection, and for the reciprocity that obtains between self-awareness and other-awareness (Ramachandran 2009).[14]

Conclusion

Aristotle is the leading philosopher of the manifest image of the world. In his hylomorphic account, the form is the actuality of an object, since its occurrence explains why this particular quantity of matter comes to

[14] Nevertheless Ramachandran noticed that: "Neurological conditions have shown that the self is not the monolithic entity it believes itself to be. It seems to consist of many components each of which can be studied individually, and the notion of one unitary self may well be an illusion. (But if so we need to ask how the illusion arises; was it an adaptation acquired through natural selection?)" (Ramachandran 2009).

be that object, in contrast to some other kind of object. Likewise, then, the occurrence of the soul explains why this matter is the living matter of an individual subject, a human being, as opposed to some other kind of object. What is significant here is that this way of construing the integral connection between the soul and the body as a special case of form-matter relations treats reference to the soul as an integral, inseparable part of any complete explanation of a living being, be it in general terms or at the more particular level of explanations of human beings.

Aristotle emphasises the importance of the soul in explanations of living beings. Concurrently with this, he considers claims about the separability of the soul from the body to be unjustified, and makes first of all an appeal to formal causation. Aristotle's hylomorphism therefore stands somewhere in between the theories of reductive materialism and Cartesian dualism.

Returning to the issue of the scientific image, let us recall that we found that individual subjects can be treated as emergent entities playing a role in unifying emergent mental properties correlated in turn with distinct modalities of the brain. The theory of strong emergence has, as a heuristic theory, quite general predictive and explanatory power as regards the field of human thought and behaviour. The emergent individuals possess a causal power to influence their own brains, meaning that thanks to the neuroplasticity of the brain, they have the power to organize or configure it. As we saw, such individuals are unique entities emerging (in a strongly emergent sense) from their brains, and influenced by their cultural surroundings as well.

Studies of neuroplasticity lend credibility to the thesis that individuals and their mental properties are not epiphenomenal entities, but rather exert an emergent form of downward causation in order to influence their own brain functions.

One of the questions we have posed concerns the extent to which the heuristic theory of strong emergence might furnish explanations of human thinking and behaviour similar in kind to Aristotelian hylomorphic theory. According to Aristotle, the occurrence of life marks a sharp ontological break in the manifest world. (However, it is worth stressing here that apart

from offering a few remarks in passing, our topic of investigation here has not really been the role of strong-emergence theory in contemporary biology.) As regards the relation between the individual subject and its brain, we have come down on the side of the relation of strong emergence. As of today, that relation would seem to provide an adequate explanation of the connectedness observed to obtain between the manifest-image understanding of a person on the one hand, and the subject qua individuum, with its brain as the object of scientific investigation, on the other.

We have also seen that a more definitive outcome stemming from Libet's experiments is the realization that a person can simply choose to perform or not perform (i.e., veto) their volitional actions. (His own philosophical explanation of this, as we saw, relies on his theory of the unified conscious field.)

Even so, we do need further—conceptual and empirical—research to answer the question of whether the unified individual subject in the framework of the scientific image should be construed on the basis of the model of Aristotelian substance or of Hume's "bundle theory of the self"—or, for example, on the basis of the model of a unified conscious-field theory.

Cartesian dualism, supplemented with Locke's account of qualities, has become the main philosophical influence on cognitive-neuroscientific research, exercising such a role over the course of three centuries and right up to the present time. Although cognitive neuroscientists working today have largely abandoned substance dualism, they continue to ascribe mental attributes to parts of human beings, typically their brains. Such attributions, however, are logically flawed or fallacious.

The Aristotelian concept of soul is preferable to that of Cartesian mind. It should therefore be regarded as providing the proper conceptual basis for both scientific psychology and the philosophy of mind. What this means is that the theoretical framework within which most scientific psychological and neurophysiological work is being carried out is—if only implicitly—the methodological successor, in actual fact, of the manifest-image (Aristotelian) one.

Summing up, the core of our argument here for adopting the concept of soul (psyche) in place of the notion of mind has been that over the course

of time the Cartesian conception of soul came to supplant the Aristotelian one in a manner that can be described as *pars pro toto*: the thinking soul was itself identified with the entirety of Aristotelian soul. As a consequence, the philosophy and psychology of consciousness, and the neurosciences, have all come in essence to amount to sciences deprived of soul in the Aristotelian sense. The seeming possibility of explaining the body's functions, and of conceiving of the latter generally, without invoking the concept of soul (i.e., treating it as just a mechanical system), doubtless helped foster such an outcome.

References

Acharya, Sourya, and Samarth Shukla. 2012. "Mirror Neurons: Enigma of the Metaphysical Modular Brain." *Journal of Natural Science, Biology and Medicine* 3 (2):118–24. Accessed October 8, 2016. doi: 10.4103/0976-9668.101878.

Aristotle. 1986. *On the Soul*. Translated by W.S. Hett. Cambridge, MA: Harvard University Press; London: William Heinemann Ltd. https://www.scribd.com/document/82577288/Aristotle-On-the-Soul-Translated-by-Hett.

Aristotle. 1993. *Posterior Analytics*. Second Edition. Translated with a commentary by Jonathan Barnes. Oxford: Clarendon Press.

Aristotle. 2009. *Physics*. Book III. Translated by Robert P. Hardie and Russell K. Gaye. Accessed October 8, 2016. http://classics.mit.edu/Aristotle/physics.3.iii.html.

Aristotle. 2017. *Metaphysics*. Accessed October 8, 2016. doi: 10.4159/DLCL.aristotle-metaphysics.1933.

Bremer, Józef. 2005. *Jak to jest być świadomym. Analityczne teorie umysłu a problem neuronalnych podstaw świadomości*. Warszawa: IFiS PAN.

Bremer, Józef. 2007. *Rekategorisierung statt Reduktion.* Wilfrid Sellars' Philosophie des Geistes. Göttingen: Vandenhoeck & Ruprecht.

Bremer, Józef. 2010. *Wprowadzenie do filozofii umysłu*. Kraków: WAM.

Bremer, Józef. 2013. *Czy wolna wola jest wolna? Kompatybilizm na tle badań interdyscyplinarnych*. Kraków: WAM.

Bremer, Józef. 2014. "René Descartes: Dancing and Mustering Substances." *The Ignatianum Philosophical Yearbook* 20 (1):7–26. Accessed October 8, 2016. doi: http://dx.doi.org/10.5281/zenodo.44462.

Bremer, Józef. 2015. "Przyczynowość skierowana ku dołowi i jej rozumienie w biologii." *Poznańskie Studia z Filozofii Nauki* 42 (1):93–115.

Bunnin, Nicholas, and Jiyuan Yu, eds. 2014. *The Blackwell Dictionary of Western Philosophy*. Malden, MA: Blackwell Publishing.

Carvalho, Diana, Silmar Teixeira, Marina Lucas, Ti-Fei Yuan, Fernanda Chaves, Caroline Peressutti, Sergio Machado, Juliana Bittencourt, Manuel Menéndez González, Antonio E Nardi, Bruna Velasques, Mauricio Cagy, Roberto Piedade, Pedro Ribeiro, and Oscar Arias-Carrión. 2013. "The Mirror Neuron System in Post-stroke Rehabilitation." *International Archives of Medicine* 41 (6). Accessed October 8, 2016. doi: 10.1186/1755-7682-6-41.

Chalmers, David J. 2006. "Strong and Weak Emergence." In *The Re-emergence of Emergence*, edited by Philip Clayton and Paul Davies, 244–55, Oxford: Oxford University Press.

Cheung, Katharine L., Eugene Tunik, Sergei V. Adamovich, and Lara A. Boyd. 2014. "Neuroplasticity and Virtual Reality." In *Virtual Reality for Physical and Motor Rehabilitation*, edited by Patrice L.T. Weiss, Emily A. Keshner, and Mindy F. Levin, 5–24. New York: Springer Science+Business Media.

Damasio, Antonio. 2010. *Self Comes to Mind: Constructing the Conscious Brain*. New York: Pantheon Books.

de Vries, Sjoerd, and Teo Mulder. 2007. "Motor Imagery and Stroke Rehabilitation: A Critical Discussion." *Journal of Rehabilitation Medicine* 39:5–13. Accessed October 8, 2016. doi:10.2340/16501977-0020.

Deacon, Terrence, and Spyridon Koutroufinis. 2014. "Complexity and Dynamical Depth." *Information* 5:404–23. Accessed October 8, 2016. doi:10.3390/info5030404.

Gazzaniga, Michael. 2011. *Who's in Charge?: Free Will and the Science of the Brain*. New York: Harper Collins.

Gazzaniga, Michael. 2013. "Understanding Layers: From Neuroscience to Human Responsibility." In *Neurosciences and the Human Person: New Perspectives on Human Activities* (Scripta Varia 121), edited by Antonio Battro, Stanislas Dehaene, Marcelo Sánchez Sorondo, and Wolf Singer, 1–14. Vatican City: Pontifical Academy of Sciences. Accessed October 8, 2016. http://www.casinapioiv.va/content/dam/accademia/pdf/sv121/sv121-gazzaniga.pdf.

Goethe, Johann Wolfgang von. 2003. *Faust*. Translated by Anthony S. Kline. Accessed October 8, 2016. http://www.iowagrandmaster.org/Books%20in%20pdf/Faust.pdf.

Goodman Lenn E., and D. Gregory Caramenico. 2010. *Coming to Mind: The Soul and Its Body*. Chicago: University of Chicago Press.

Hacker, Peter M.S., and Max R. Bennett. 2003. *Philosophical Foundations of Neuroscience*. Malden, MA: Blackwell.

Hodgson, David. 2012. "Identifying and Reconciling Two Images of 'Man'." *Humana. Mente – Journal of Philosophical Studies* 21:45–56.

Kim, Jaegwon. 2006. "The Exclusion Argument." In Jaegwon Kim. *Philosophy of Mind*, 195–96. Boulder, CO: Westview Press.

Krohs, Ulrich, and Georg Töpfer, eds. 2005. *Philosophie der Biologie*. Frankfurt am Main: Suhrkamp.

Libet, Benjamin. 2003. "Can Consciousness Experience Affect Brain Activity", *Journal of Consciousness Studies*, 10: 24–28.

Llaca-Rodríguez, Victoriano, and Santiago Ramón y Cajal. 2001. "Genio inmortal y pensador universal", *Cirugia y Cirujanos* 69 (4):201–206.

Mateer, Catherine A., and Nicholas M. Bogod. 2003. "Cognitive, Behavioral, and Selected Pharmacologic Interventions in Rehabilitation after Acquired Brain Injury." In *Neuropsychiatry*, edited by Randolph B. Schiffer, Stephen M. Rao, and Barry S. Fogel, 165–71. Philadelphia, PA: Lippincott Williams & Wilkins.

Merabet, Lotfi B., and Alvaro Pascual-Leone. 2010. "Neural Reorganization Following Sensory Loss: The Opportunity of Change." *Nature Reviews. Neuroscience* 11:44–52. Accessed October 8, 2016. doi:10.1038/nrn2758.

Nahum, Mor, Hyunkyu Lee, and Michael M. Merzenich. 2013. "Principles of Neuroplasticity-Based Rehabilitation." In *Changing Brains: Applying Brain Plasticity to Advance and Recover Human Ability*, edited by Michael M. Merzenich, Mor Nahum, and Tom M. Van Vleet, 141–73, Amsterdam: Elsevier.

Page, Stephen J., Peter Levine, Sue A. Sisto, and Mark V. Johnston. 2001. "Mental Practice Combined With Physical Practice for Upper-Limb Motor Deficit in Subacute Stroke." *Physical Therapy* 81 (8):1455–62. Accessed October 8, 2016. doi.org/10.1093/ptj/81.8.1455.

Ramachandran, Vilayanur S. 2009. "Self-Awareness: The Last Frontier." Edge Foundation web essay. Accessed October 8, 2016. http://www.edge.org/3rd_culture/rama08/rama08_index.html.

Rosenberg, Jay A., and Wilfrid Sellars. 2007. *Fusing the Images*, Oxford: Oxford University Press.

Schultze-Kraft, Matthias, Daniel Birman, Marco Rusconi, Carsten Allefeld, Kai Görgen, Sven Dähne, Benjamin Blankertz, and John-Dylan Haynes. 2016. "The Point of No Return in Vetoing Self-initiated Movements." *Proceedings of the National Academy of Sciences of the United States of America* 113 (4):1080–5. Accessed October 8, 2016. doi: 10.1073/pnas.1513569112.

Schwartz, Jeffrey M., and Sharon Begley. 2002. *The Mind and the Brain: Neuroplasticity and the Power of Mental Force*. New York: Regan Books; Harper Collins Publishers.

Schwartz, Jeffrey M., and Rebecca Gladding. 2012. *You Are Not Your Brain: The 4-Step Solution for Changing Bad Habits, Ending Unhealthy Thinking, and Taking Control of Your Life*. New York: Avery Publishing Group.

Seager, William. 2012. *Natural Fabrications: Science, Emergence and Consciousness*. Berlin: Springer-Verlag.

Sellars, Wilfrid. 1949. "Aristotelian Philosophies of Mind." In *Philosophy for The Future: The Quest of Modern Materialism*, edited by Roy Wood Sellars, V.J. McGill, and Marvin Farber, 544–70. New York: The Macmillan Co.

Sellars, Wilfrid. 1963. "Philosophy and the Scientific Image of Man." In Wilfrid Sellars. *Empiricism and the Philosophy of Mind*, 1–40, London: Routledge & Kegan Paul Ltd.

Sellars, Wilfrid. 1974. "Metaphysics and the Concept of a Person." In Wilfrid Sellars. *Essays in Philosophy and Its History*, 214–244. Dordrecht: D. Reidel Publishing Company.

Sellars, Wilfrid. 1975. "The Structure of Knowledge." In *Action, Knowledge and Reality: Studies in Honor of Wilfrid Sellars*, edited by Hector-Neri Castaneda, 259–347. Indianapolis, IN: Bobbs-Merrill.

Sellars, Wilfrid. 1992. "Persons." In Wilfrid Sellars. *Science and Metaphysics: Variations on Kantian Themes*, 170–1. Atascadero, CA: Ridgeview Publishing Company.

Sen, Amartya. 2000. *Development as Freedom*. New York: Anchor Books.

Sperry, Roger. 1976. "Mental Phenomena as Causal Determinants in Brain Function." In *Consciousness and the Brain: A Scientific and Philosophical Inquiry*, edited by Gordon G. Globus, Grover Maxwell, and Irwin Savodnik, 163–77. New York: Plenum Press.

Sperry, Roger. 1987. "Consciousness and Causality." In *The Oxford Companion to the Mind*, edited by Richard L. Gregory, 164–6. Oxford: Oxford University Press.

Stephan, Achim. 2002. "Emergentism, Irreducibility, and Downward Causation." *Grazer Philosophische Studien* 65:77–93.

Strawson, Peter F. 1959. *Individuals*. London: Methuen & Co Ltd.

Tateo, Luca. 2015. "Just an Illusion? Imagination as Higher Mental Function." *Journal of Psychology and Psychotherapy* 5 (6):1–6. Accessed October 8, 2016. doi.org/10.4172/2161-0487.1000216.

Witt, Charlotte. 1992. "Dialectic, Motion, and Perception: De Anima, Book I." In *Essays on Aristotle's De Anima*, edited by Martha C. Nussbaum and Amélie Oksenberg Rorty, 169–84. Oxford: Clarendon Press.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. Translated by Gertrude Elisabeth M. Anscombe. Oxford: Blackwell.