

Işık Sarıhan

**Philosophical Puzzles Evade Empirical Evidence: Some Thoughts and Clarifications
Regarding the Relation between Brain Sciences and Philosophy of Mind**

Penultimate draft. Final version published in Leefman, J & Hildt E. (eds.) (2017) *Human Sciences after the Decade of the Brain*. Elsevier, San Diego, CA, 14-23.

Abstract:

This chapter analyzes the relation between brain sciences and philosophy of mind, in order to clarify in what ways philosophy can contribute to neuroscience and neuroscience can contribute to philosophy. Especially since the 1980s and the emergence of “neurophilosophy”, more and more philosophers have been bringing home morals from neuroscience to settle philosophical issues. I mention examples from the problem of consciousness, philosophy of perception and the problem of free will, and I argue that such attempts are not successful in trying to settle questions like whether psychology can be reduced to neuroscience, whether we see the external world directly in perception, or whether we have free will. The failure results from an ability of the philosophical questions to evade the data. What makes these questions persisting philosophical questions is precisely that there is no way to settle them through empirical evidence, as they are conceptual questions and their solution lies in conceptual analysis.

Keywords: Metaphilosophy, philosophical methodology, neurophilosophy, philosophy of psychology, philosophy of perception, free will

In this chapter, I chart the general territory of interdisciplinary interaction between analytic philosophy and brain sciences and the relevance the two fields have for each other. I differentiate two main styles of interaction, from philosophy-to-neuroscience and from neuroscience-to-philosophy. The first type of interaction is motivated by the philosophers' interest in the issues that arise within neuroscience, the findings of the field and the claims made by neuroscientists in the context of the philosophers' function as a conceptual and logical corrector or aide towards a more accurate science. The other aspect concerns the relevance of the findings of neuroscience for the resolution of philosophical debates. I will make some brief remarks about the first, and then investigate more closely the second type of interaction. I will not discuss the controversial field called "experimental philosophy".

Especially since the 1980s and the emergence of "neurophilosophy", more and more philosophers have been closely following the findings of neuroscience and other related sciences, and bringing home morals for philosophical questions. I will discuss whether these attempts succeed or not in working towards a resolution of philosophical debates. I will mention two cases, one regarding a puzzle in philosophy of perception and one regarding the mind-body problem, and I will argue that such attempts are not successful in trying to answer questions like whether we see the external world directly in perception or whether psychology can be reduced to neuroscience.

This failure is not due to any problems with the neuroscientific data itself, but it results from an ability of the philosophical questions to evade the data, that is, the data mentioned fails to settle the

philosophical debates conclusively. This has a reason: If those philosophical questions could be settled by empirical evidence, they wouldn't be philosophical questions in the first place, they could be reframed as scientific, empirical questions. What makes these questions persisting philosophical questions is precisely that there is no way to settle them through empirical evidence. Rather, they are conceptual questions, some of them also with phenomenological aspects, and their solution lies in conceptual analysis and/or phenomenological methods.

1. Philosophy-to-Science: Ethics, logic, terminology and clarity

There are many ways philosophy functions as an aide and an inspector for science. For continental philosophy, this is generally by highlighting and discussing various normative, political and societal aspects of scientific theories, terminologies and processes; but this article is about analytic philosophy, which I will refer simply as “philosophy” from now on. I will very briefly identify some functions of philosophy in relation to science.

One contribution of analytic philosophy to sciences is through introducing new conceptual and terminological tools to a field (or the criticism of the introduction of such tools). For instance, the use of the term “representation” in cognitive neuroscience comes through a long philosophical tradition, a famous example promoting the use of such a conceptual scheme being Fodor's seminal work (1975). This is of course something that is also often done by people who are not necessarily philosophers by profession but who are working on the theoretical levels in a scientific field. Another function common to both philosophers and scientists is asking questions that can inspire scientific work. Indeed, the philosopher and the theoretical scientist is hard to distinguish at this level. The philosophizing aspect of the philosopher comes into the picture where the contribution

to be carried out by the philosopher involves the kind of conceptual analysis and logical argumentation typically found in analytic circles.

One thing a philosopher can do, qua philosopher, is to criticize a certain idea that appears in a scientific context on the grounds that it doesn't logically follow from the evidence, or to criticize a terminology on the grounds that concepts deployed are problematic. To give an example, Bennet and Hacker criticize the conceptual background of neuroscience from a Wittgensteinian metaphilosophical perspective, arguing that many terms used in neuroscience are confused, in the sense that they rest on misunderstandings of terms in everyday vocabulary (Bennet and Hacker, 2003). Another prominent example of this type of contribution is the debate that sparked by the empirical findings of Libet on the basis of which some claimed that we don't have free will because before we are aware of a mental act of decision-making, a certain unconscious process which factors into which decision is to be made is already going on in our brains (Wegner, 2002). This conclusion was challenged by philosophers in many ways, for example, by claiming that it relies on a misunderstanding of the term "free will": When one analyzes the concept, it is claimed, one sees that in order to freely perform actions one doesn't need to be aware of a mental act of "deciding" or "willing"; in order for an action to be free other criteria are sufficient, such as the action's being in line with one's beliefs and desires, the decision not being taken under duress, etc. (See O'Connor 2010 for a review of various analyses of free will; also see Dennett 2003, ch. 8, for a criticism of Wegner.)

We should recognize that many such debates around science are actually spillovers from philosophers' debates. For instance, "qualia" is a term that has originally appeared in philosophical literature, and started to appear in certain scientific work especially after the 90s (see for instance Ramachandran and Hirstein 1997, Edelman and Tononi 2000). But many philosophers claim that

the concept of qualia is confused and attempts to refer to something that doesn't exist, if it is attempting to refer to an internal quality, and according to this view scientists who are trying to solve the problem of qualia are dealing with a pseudo-problem. (Representationalists like Byrne (2006) or Wittgensteinians like Bennett and Hacker (2003) are examples.) This is nothing other than the philosophers' debate carrying itself over to the battleground of science. Similarly, going back to the Libet example above, Dennett (2003) has claimed that the confusion regarding free will based on Libet's experiments has a long philosophical background that goes back at least to Descartes.

One last function of philosophy that should be mentioned is the role played by ethics in the context of science. Ethics of neuroscience is a relatively well-established field, involved in various important worldly matters such as neuromarketing, neuro-enhancement and animal experimentation.

2. Science-to-Philosophy: Can scientific data settle philosophical debates?

Now we will look at the other side of the interaction, taking empirical data and trying to come to conclusions regarding the traditional questions of philosophy. By going through two cases, one regarding the question of whether we see the outside world directly or not, and the other regarding the question of whether the mind can be reduced to the brain, I will try to show this direction of interaction is not very fruitful, for systematic reasons due to the nature of philosophical questions. There can of course be other ways of science-to-philosophy direction of intellectual contribution which I will not talk about, say, some new empirical data providing a good case example to discuss something in a philosophical context or triggering novel philosophical questions. An empirical

finding can also have some relevance for ethics, for instance, if it turns out that a creature feels pain and pleasure we are obliged to behave differently towards that creature. It should be remembered, however, that we have these obligations already in the context of a background ethical code, itself a matter of philosophical debate untouched by empirical evidence, like not causing unnecessary harm or not terminating otherwise pleasurable lives.

2.1. Case one: Philosophy of perception

One philosophical question that persisted for centuries is whether we see the external world directly or not, whether the things we seem to be directly acquainted with in perception are objects and properties in the external world or whether they are internal or intermediary items in some sense. These intermediary items have went with various names, most famously “sense-data” or sometimes “sensations”. More recently, some have adopted the term “qualia” for this terminological purpose (Wright, 2008), even though more often the term has been deployed to mean something slightly different.

Many philosophers and scientists have insisted that the findings of many empirical fields, particularly psychophysics and neuropsychology of conscious experience, has given us proof that “indirect realism”, the philosophical view that we don’t perceive the external world directly, is the correct view (Revonsuo, 2006, also see science-based articles in Wright, 2008). According to this approach, this centuries-old philosophical question has been waiting for the emergence of scientific evidence to be settled. However, unfortunately, the evidence doesn’t settle the issue, and even if it did, much simpler evidence could be just enough, without extensive modern-day research.

The science-based approach to settle this question relies mainly on the fact that every type of experience one can have depends on a certain internal, neural state, which seems to be enough to have that type of experience without an external corresponding object, as it is the case in illusions, dreams and hallucinations. One can see a sunset in a dream, one can see an object as blue even if it is some other color, or still have the lingering phantom feeling of an amputated arm. Since the hallucinatory cases and the externally-caused veridical counterparts are subjectively the same experiences, and since we are both aware of something in both cases, then, just like we are not directly aware of an external entity in the illusory cases, then we are not directly aware of the external world in the veridical cases either. Rather, what we are directly aware of are models or internal images constructed by the brain. This idea is also supplemented by the finding that there is a lot of constructive work going on in one's brain during ordinary perception, say, the filling-in of the blind spot or various other "inferences" the perceptual system makes to construct an image of the world.

This old philosophical argument that predates contemporary science is known as the "argument from error" (also "argument from hallucination" or "argument from illusion") in the philosophical literature. Note that if the argument could settle the debate, it could have settled it a long time ago without help from rigorous empirical science of today, since hallucinations, illusions and dreams have long been well-known phenomena. One doesn't need contemporary neuroscience to know that one can have the experience of seeming to see a particular object or particular quality without there being a corresponding thing in the environment. So, the first point is that, even if the empirical data could settle the issue, hard science wouldn't be necessary.

The second point is, the empirical evidence cannot settle the issue because the conclusion of the argument, that we do not see the world directly, rests on a particular understanding of certain terms

like “seeing”. A common objection to the argument from error has been that when we hallucinate something, we don’t really see anything, rather, we “seem to see” something, we have a hallucinatory experience of seeing something with our eyes. So one cannot generalize from hallucination to veridical perception and say that what we see in both cases is the same and therefore in veridical perception we don’t see the external world directly. In hallucination we are aware of something for sure, the “content” of our experience, but this doesn’t make it impossible for us to be in direct empirical touch with the external world. In hallucination, the content doesn’t match the outside world, in veridical perception, it matches, and via “having” this content (as opposed to “seeing” or “perceiving” this content) we are aware of the external world directly without being aware of anything else. It is indeed the case that the brain has some constructive role in everyday perception, and the sensory content regarding what is in front of us provided by this constructive processes partly distort our vision of what is out there, but still there is no conceptual obstacle to say that for what corresponds out there to the undistorted bit of the content, we see those aspects of the world directly. Think of a visual state which is partly illusory, you experience a scene which represents the world as mainly as it is but which includes the illusory appearance of a bent stick in water. Some would like to analyze this perceptual state as the brain constructing an image of the world that is partly incorrect, and the immediate object of our awareness is this internal image. But the brain’s construction of an image, a model, or an appearance of the outside world need not be analyzed in this epistemic way. One can as well say that as a result of some external and internal processes, the world appears to us in a certain way, and if this appearance is truthful, we see the world directly, and the misleading representation of the bent-stick-in-water shouldn’t be understood as “seeing” an image of a bent stick, but failing to see a particular spatial property of the stick because of having a misleading visual representation of the world.

This is one among various responses given to the argument from hallucination. I am not interested in logically proving it here, but to demonstrate that one's reliance on empirical data to argue for indirect realism rests on a certain understanding of certain concepts like seeing, awareness, perception, etc., and this conceptual debate cannot be settled by empirical evidence. And if it could, we wouldn't need hard science, since the mere existence of illusions and hallucinations could prove this point.

2.2. Case two: The mind-body problem

The mind-body problem is the problem of understanding what the relation between the mind and body is, or more precisely, whether mental phenomena are a subset of physical phenomena or not. There are many philosophical positions associated with this problem, substance dualism ("mind and body are two different substances"), property dualism ("there is only one, physical substance, but mental properties of subjects cannot be reduced to their physical properties"), and physicalist reductionism ("mental properties can be identified with, or can be spelled out in terms of, physical properties"), among other positions.

Some philosophers in the past decades have argued that modern neuroscience has already given us an answer to this question: Mental states are nothing other than neural states, and we can talk of mental phenomena through physical vocabulary without any loss of meaning or reference. Founders of what is called "neurophilosophy", Patricia and Paul Churchland, have been among the most famous advocates of this position, even though their views oscillated between reductionism and eliminativism, the latter view being that mentality (or certain aspects of it) is a pre-scientific

construct that will have no place in the scientific understanding of the world once we have a fully developed neuroscience (See for instance Churchland, 1988).

This brand of reductionism relies heavily on the findings of neuroscience, often on quite detailed empirical data about the relation between certain psychological states and brain states or how neural processes generate behavior. Well, almost all brands of reductionism have some reference or other to the brain science, but they are often uninterested in detailed data. Rather, they simply point to the general scientific consensus that there is a very direct relation between mental and neural states, but this relation by itself does not play an important role in their arguments, for these arguments generally rely on a conceptual analysis of mental states to see if mental states could be reduced to *any* physical states to begin with, and if such reduction is possible, modern day science tells us which physical states are the reduction base, which turns out to be neural or bodily states. On the other hand, the kind of reductionism we are interested in here takes it that *science has proven* mind-brain identity.

The scientific evidence, however, does not settle the philosophical question. No matter what detailed and direct mapping we establish between mental and neural states, there are so many options that remain on the table before we proclaim that we have reduced mental processes to brain processes. The first obstacle is that correlation does not mean identity, and the reductionist should answer arguments to the effect that the relation is better explained as a correlation. The most important and pervasive argument I will mention here is the argument that, to put it roughly, the reduction of mental phenomena to physical phenomena does not make sense, on conceptual or logical grounds. A phenomenon described by a physical description like “such and such connectivity and firing in this and that brain area” simply cannot be identical to “feeling pain” or “thinking about Budapest” or “having a visual experience of a yellow lemon” as these mental

phenomena described have certain characteristics that the physical description does not capture, even though anti-reductionist philosophers disagree on what these characteristics are. Some say it is some subjective aspect, the “feel” or “what-it-is-likeness” associated with the mental state, some say it is “intentionality”, the property of “having a content” or “being directed to an object”, such as the object experienced or thought about. When we are given a description like “such and such neural firing”, we cannot infer whether this state is an experience of color green or color red or something else. This disparity is termed “the explanatory gap”: A successful reduction, it is argued, should tell us not only an identity based on an observation of a correlation between the occurrence of mental and neural states, but also make us understand how is it possible that such and such mental firing could constitute experiencing the color red (Levine, 2001). Another very different anti-reductionist argument, a behaviorist one, is the argument that the mind cannot be the brain because mental terms do not strictly refer to states or processes, but dispositions to behave, and it doesn’t make sense to identify a disposition with a state or process (Ryle, 1949)

Philosophers have answers to these challenges of course. Some claim that identities don’t require an explanation (Block and Stalnaker, 1999), some try to analyze mental phenomena into physical or “topic-neutral” vocabulary (vocabulary that is common both to physical and mental terminology, such as the vocabulary of basic ontological phenomena like causation; the approach was popularized by Smart (1959) and many causal-informational analyses of mental phenomena can be considered as a continuation of this strategy). To give a more specific example, some philosophers like Millikan analyze mental states as natural indicators, “thinking of X” is being in a state which has the function of indicating the presence of X. So in case of humans, it is a brain state that has acquired the biological function of signaling the presence of X through the organism’s interaction with its environment (Millikan, 1984). Again, it is not important for us here whether

this analysis is correct or not. But if this analysis is correct, then there is no obstacle for mental states to be reduced to neural states ---as instances at least, if not as scientific or functional kinds, since different types of neural states can realize the same mental state--- and one doesn't need to look at detailed empirical data to claim that reduction is possible, given that every mental state will correlate with some neural state. From then on, if Millikan's analysis is accepted, whether mental state M is identical to neural state N is philosophically uninteresting in the context of the mind-body problem, that is, further empirical evidence do not add anything to the solution of the philosophical problem.

Conclusion

The two examples above have shown us that philosophical puzzles cannot be settled by empirical evidence, and it is not due to any specific features of these two cases, it is something that results from the nature of philosophical questions. Philosophical questions are often classified as questions of a logical or conceptual nature. But let's avoid here the task of defining what a philosophical question is, it is enough to mention a negative aspect: Philosophical questions are not empirical questions. If they could be conclusively settled down by empirical evidence, then they would be turned into scientific questions and taken from the hands of philosophers. This is arguably what happened with some philosophical questions in the past, questions that stopped being philosophical questions and turned into questions of physics or biology, such as some questions regarding spatio-temporal behavior of objects, questions regarding basic elements, some questions related to vision, and issues about the origins and diversity of life forms. (Think of ancient Greek theories of objects moving towards their "natural place" or the philosophical theory of vision that postulated beams emitted by the eyes.) Answering questions like whether we see the world directly or whether

mentality can be reduced to physics requires first of all an agreement on what “seeing the world directly” means or what we mean by “mentality” or “reduction”. And when they are settled, we still won’t be able to say that the “philosophical question” can be answered empirically, because the philosophical part of the question would already be solved.

Another point about this non-empirical nature of philosophical questions is that their solution is of little or no practical concern. If we cannot settle a question empirically, that is, by intervening with the world and then observing the outcomes, then the settlement of the question won’t give us any advantage of intervening with the world to bring about outcomes either. Of course if a philosophical question has an ethical or normative aspect, then one’s converging on one answer or the other has consequences for one’s personal behavior or how we conduct science. But otherwise, the resolution of the mind-body problem or the philosophical problems of perception don’t seem to add anything to our ability in scientific contexts, in terms of, say, prediction or engineering. No matter the mental properties are identical to neural properties or are just irreducible and systematically correlated with neural properties, we still get the same results in neuroimaging or behavioral neuroscience. No matter we see the world directly or not, we get the same results in psychophysics. Still, one can expect a scientist to be careful about what background philosophical view she takes, if she takes any at all, while using a set of terms and concepts in relation to the scientific work. This means either putting the philosophical questions aside, or, when making big claims, being aware of rich and complicated debates going on in the contemporary philosophical background surrounding such issues. Well, the very act of looking into the brain and acquiring the data itself rests on certain background philosophical assumptions, but many are unimportant in the particular context of the interaction of neuroscience and philosophy of mind. Instead, they are related to general philosophical issues surrounding science and observation, and a scientist cannot be expected to

worry about all such philosophical problems before taking up some work. But after the empirical part is done, when one starts making claims like there is no free will, the mind is the brain or we don't see the external world directly, one has stepped into the domain of philosophy of mind.

References

Bennett, M.R. & Hacker P.M.S. (2003) *Philosophical Foundations of Neuroscience*. New Jersey: Wiley-Blackwell.

Block, N. & Stalnaker, R. (1999) "Conceptual analysis, dualism, and the explanatory gap", *Philosophical Review*, 108(1):1–46.

Byrne, A. (2006) "Color and the mind-body problem", *Dialectica* Vol. 60, No. 3: 223–244.

Dennett, D.C. (2003) *Freedom Evolves*. New York: Viking Press.

Churchland, P.M. (1988) *Matter and Consciousness*. Cambridge, MA: MIT Press.

Fodor, J. A. (1975) *The Language of Thought*. Cambridge, Massachusetts: Harvard University Press.

Edelman, G.M. & Tononi, G. (2000) *A Universe of Consciousness: How Matter Becomes Imagination*. New York: Basic Books.

Levine, J. (2001) *Purple Haze: The Puzzle of Consciousness*. Oxford University Press.

Millikan, R. (1984) *Language, Thought and Other Biological Categories*. Cambridge MA: MIT Press.

O'Connor, T. (2014) "Free will", *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2014/entries/freewill/>

Ramachandran, V.S. & Hirstein, W. (1997) "Three laws of qualia: What neurology tells us about the biological functions of consciousness, qualia and the self", *Journal of Consciousness Studies*, 4, No. 5-6: 429–58.

Revonsuo, A. (2006) *Inner Presence: Consciousness as a Biological Phenomenon*. Cambridge, MA: MIT Press.

Ryle, G. (1949) *The Concept of Mind*. London: Hutchinson.

Smart, J.J.C. (1959) "Sensations and brain processes", *Philosophical Review*, 68: 141–156.

Wegner, D. (2002) *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.

Wright, E. (ed.) (2008) *The Case for Qualia*. Cambridge MA: MIT Press.