

Explanatory Gap and Mental Causation

Reena Cheruvalath

RESUMEN

Desde su origen en los debates filosóficos, el concepto de conciencia ha permanecido como algo problemático. La naturaleza de la conciencia genera siempre una «brecha explicativa» entre ella misma y los estados físicos. John Searle trata de cerrarla con su propuesta, conocida como «naturalismo biológico». Su teoría de la causación mental descansa en el principio de sobrevenida y la diferenciación de niveles. Esto supone una mezcla de fisicalismo no-reductivo, misterianismo, emergentismo, dualismo, epifenomenalismo modificado y realismo, incluso si el propio Searle rechaza todas estas etiquetas. En consecuencia esto abre la puerta a una brecha insuperable tanto en el nivel fisiológico como psicológico.

ABSTRACT

Ever since its origins in the philosophical debate the concept «consciousness» has remained as problematic. The nature of consciousness always invites an «explanatory gap» between itself and the physical states. John Searle tries to fasten it through his view, which is known as «Biological Naturalism». His theory of mental causation is supported by a supervenience principle and levelism. It is a mixture of non-reductive physicalism, mysterianism, emergentism, dualism, modified epiphenomenalism and realism, even if he rejects all those labels. Accordingly, it opens the door for an unbridgeable gap both at the physiological and psychological level.

Ever since its origins in the philosophical debate the concept «consciousness» has remained as problematic. For, when we discuss it in relation to the body, the nature of consciousness always invites an «explanatory gap» between itself and the physical states. It means that there is no adequate explanations for questions like: «How is it possible for the mind to cause a change in a material body?» «How does a mental event manage to initiate, or insert itself into, a causal chain of physical events?» «How is it possible that a chain of physical and biological events and processes terminates in a full-blown conscious experience?» and «How is it possible for sensory experience to arise out of the electrochemical processes in the grey matter of the brain?» John Searle tries to fasten it through his view, which is known as «Biological Naturalism». He presents a biological naturalist's account of mental causa-

tion to formulate a tie up between mental states and physical states and thereby to close the gap between brain and consciousness.

I. THE PROBLEM OF THE EXPLANATORY GAP

The explanatory gap can be described as occurring between either two sets of substances or two sets of properties, or else, it might be about the relation between these sets of properties or, still again, it might even be the relation between two events. It arises when we lack any bridging principle between some mental facts and the other, and also between mental facts and physical facts. This is changed into a biconditional law and thenceforward into identity which underwent yet another modification in terms of a theory of mental causation. The term «explanatory gap» was first coined by Levine (1983) to show that consciousness cannot be explained simply by telling that it accompanies physical facts. Levine argues that the gap is an epistemological one that is compatible with the thesis that facts about consciousness supervene on the physical facts [Levine (1983), pp. 354-361].

Mainly, there are three different arguments, which one may reasonably seize on the issue.

- a) The explanatory gap is unbridgeable and no purely physical explanation is possible.
- b) The gap may one day be bridged, but we currently lack the concept to bring the subjective and objective perspective together.
- c) There is no such explanatory gap at all.

The first argument establishes the fact that mental states are different from physical states. It has been supported by Karl Popper (1994) and Chalmers (1996). According to the second argument phenomenal states are physical, but we currently have no clear conception as to how they could be [Nagel (1974), pp. 435-456]. Besides, the recent revolution in the study of DNA has shown that experiences and feelings are as much part of physical, natural world as life, photosynthesis or lightening [Crick and Koch (1990), pp. 263-275]. What we call consciousness seems to be several processes, which we are lumping together in one word. And it is possible we could dissect these different processes and map them in different brain structures. The subjective consciousness or «qualia» remains as problematic however; one day we may achieve a mature enough understanding of what it is [Ramachandran (2006), pp. 5-6]. Indeed, there are also supporters of the final argument. As

the philosopher Dennett has emphasized, the explanatory gap is due to our difficulty in understanding. Churchlands' connectionist cognitive model adds that a neuro-computational perspective will be adequate to close the explanatory gap. They are also known as eliminativists. Eliminativists generally agree that reduction and supervenience are the norm in physical sciences; but such reductions are not available for mental phenomena. Thus, the objects postulated by «Folk psychology» do not exist. The reductionists who support this view are hardheaded to argue that mental states and physical states are reducible to one another and underscore that mental states are physical states or simply brain states. Thus, mental causation is a species of physical causation.

The anti-reductionists challenge this principle so as to argue that there is something over and above the physical world, which cannot be causally explainable. Consequently, it seems that the controversy about closing the explanatory gap rages mainly between reductionists and anti-reductionists. Within anti-reductionism, there are physicalists, called non-reductive physicalists, who hold that there is an explanatory gap in psychology, but that its roots lie in an inability of minds like ours to represent the explanatory relations between body and mind. In addition to this, mental properties supervene upon physical properties, but this supervenience is abidingly epistemically opaque to creatures with our particular representational capacities [Belot and Earman (1997), pp. 147-182]. This kind of non-reductionist view has served as an influential philosophical foundation to cognitive science, which forms an autonomous and irreducible science with its own distinctive vocabulary and methodology and it is not answerable to the methodological or explanatory constraints of the more fundamental sciences, such as physics and biology [Kim (1996), pp. 211-213].

As we understand it, the two master arguments, which provide the rallying point between the robust reductionists (e.g. Kim) as well as robust anti-reductionists (e.g. Ned Block) for achieving the explanatory potential are the «supervenience argument» and the «overdetermination argument». The «supervenience principle» is understood to affirm a relation of dependence or determination between the mental and the physical. In other words, our psychological character is wholly determined by our physical/biological nature. This is often read to imply a dependency thesis of this kind; for it says that once the physical nature of a thing is completely fixed, that fixes its mentality in every detail. The idea is variously used to prove the ontological primacy of consciousness (subjectivity), realism about mental states, as well as reducing the mental to the physical. So also the theory of mental causation is worked out within the theory of intentionality or within a theory of the physical world. The second argument, i.e., the overdetermination argument, affirms that actuality is made up of all the things identified by physics and anything, which is a compound of these things. So there are states as well as their microphysical constituents.

This is exactly, where our perspective on Searle receives its full impetus. Among the three viewpoints (a, b, c), which have been mentioned above, he accepts «b». He tries to bring the subjective and objective perspective together with his intentionalist theory of mental causation. Accordingly, the following question arises: «Can Searle's theory of mental causation close the explanatory gap?» This requires us to address ourselves the closely related issue of the theory of mental causation. The latter problem absorbs the former; for if Searle has a plausible theory of mental causation then he might be said to close the explanatory gap within the purview of his outlook on biological naturalism. Again, this squarely depends upon the way his intentional theory of mental causation is to be understood. That is to say, the theory of causation that depends on supervenience and a connection principle in his layered model of biological naturalism must be proved to have credentials.

II. BIOLOGICAL NATURALIST'S THEORY OF CAUSATION

Searle uses causation as a purely analytical tool in order to put forward a very provocative solution about the mind-body problem. According to him, it is possible to avoid the pitfalls of both crass materialism and idealist mysticism. In order to achieve that we must abandon the idea that causation requires physical contact between two objects; causation is not simply physical but can be mental, social, political or economic. Therefore he offers a «simple solution»: mental phenomena are «caused by» neuro-physiological processes in the brain and are themselves higher level features of the brain. His form of causation is «bottom up», whereby the behaviour of lower level elements, presumably neurons and synapses, causes the higher level or system features of consciousness and intentionality. Moreover, mental events and processes are as much part of our biological natural history as digestion, mitosis, and meiosis or enzyme secretion [Searle (1992), p. 1]. Searle suggests that as long as we continue to talk and think as if the mental and the physical were separate metaphysical realms, the relation of brain to consciousness will forever seem mysterious, and we will not have a satisfactory explanation of the relation between neuron firings and consciousness. In fact, mental states are supervenient on neuro-physiological states. He supports causal supervenience, which he distinguishes from constitutive supervenience. By causal supervenience, he means that changes in mental states are brought about by changes in an organism's underlying physiological states.

Even granting that there ought to be a point of clarification between what he calls a supervenient causation and mental causation, his theory is open to attack. Given his distinction between «causal» and «constitutive supervenience», it would still be difficult to know whether it warrants any sharp distinction so as to take it as an «independent assumption». In fact, later on

he frankly admits that his usage vacillates between the constitutive and causal supervenience [Searle (2000), p. 19]. Besides, within Searle's perspective, supervenience is used not only to provide both the causal relation between intentionality and consciousness, but also to cut off the relational link: he provides no link between brain and consciousness. This seems to generate a paradox about his theory of mental causation. It seems his supervenience principle signifies «same cause and same effect». Thus, when he places supervenience and same cause and same effect on the table, especially in the context of mental causation, without as much denying it, his theory requires to be evaluated in terms of a full-blooded theory of mental causation on the lines developed by Kim. Kim counters Searle's use of causal supervenience and demonstrates the weakness of the idea by pointing out the dilemma. The dilemma is shown in the following way:

Suppose that an instance of a mental property, M , causes another mental property, M^* , to be instantiated—an instance of «going left to right from micro to macro» as Searle calls it—. According to Searle's biological naturalism, every mental phenomenon is caused by a neurobiological phenomenon. That is to say, this instance of M^* is caused by an instance of a neural property P^* . Here, the following questions arise: where does this instance of M^* come from?, and How does M^* get instantiated on this occasion? The two answers are: (1) *ex hypothesi*, M^* was caused to be instantiated by M , and (2) according to Searle's biological naturalism, M^* was caused to be instantiated by the neural property P^* . It looks as though the instantiation of M^* is causally overdetermined.

In order to be a viable theory of causation, it must decide on the competition between the following two apparently contradictory theses: mental causation is a species of physical causation or vice versa (bottom-up causation). This is exactly, where overdetermination is introduced. If a biological property (B) causes a mental property (M), and M has causal powers to instantiate another mental property (M^*), then this might lead to the following question: Does it mean that all mental to-mental causation has two causal properties? If so it is overdetermined [Kim (2000), pp. 19-48]. The flaw of the overdetermination argument is that the argument tries to squeeze physicalism from competition between mental and physical causation. Since the causal efficacy of mind is secured by the macro image, while the causal hegemony of physics is secured by microphysics, we have not guarantee to the effect that there is competition between the two. After all, microphysics never mentions events found within the macro image [Sturgeon (1998), pp. 421-438].

Quite opposed to the above, supervenience is also used as an explicit affirmation of the ontological primacy, or priority, of the physical in relation to the mental. Accordingly, this theory opens the possibility of explaining the mental in terms of the physical. Thus, minimal physicalism can be thought of as the philosophical basis of such explanatory practices. Since this is quite

opposed to the kind of dependence thesis of supervenience used by Searle, it becomes an easy target. This is how his use of «causal supervenience» has become a target for attack in unambiguous terms. It shows that his solution is fraught with ambiguities and difficulties. Furthermore, Searle's use of supervenience requires type-type identity. Type identical neuro-physiological causes would have type-identical mentalistic effects. Accordingly, supervenience is assigned the task of inter-level relationship. Mental states are supervenient on physical states means that higher-level properties are explainable in terms of lower-level properties. Searle's levelism supports bottom-up causation. Hence, a layered model of mental causation meets its fate; it turns out to be idiosyncratic causation [Kim (2000), pp. 54-55]. Here, Kim points out that another alternative for Searle is Emergentism. Emergentism was the first systematic formulation of non-reductive physicalism as well as of the multi-layered model of the world. It consists of three doctrines. (1) All that exists in the space-time world are the basic particles recognized in physics and their aggregates, (2) When aggregates of material particles attain an appropriate level of structural complexity, genuinely novel properties emerge to characterize these structured systems. (3) Emergent properties are irreducible to, and unpredictable from, the lower-level phenomena from which they emerge [Kim (1996), pp. 226-236].

Emergentists hold that once the mental properties have emerged, these higher-level properties begin to lead a life of their own, so to speak, and manifest their powers by causally affecting lower level phenomena. This is the downward causation, the causal influence exerted by higher-level phenomena on the processes going on at a lower level. Downward causation is a fundamental commitment of Emergentism and the basic tenets of non-reductive physicalism lead to a commitment to mental-to-physical causation, a form of downward causation. In this sense, Searle cannot be called a «magical emergentist», because he prefers bottom-up causation. Likewise, emergentism subsumes property dualism, which Searle refuses to accept. For the property dualist, consciousness is a distinct, non-physical feature of the brain, but for Searle it is a state the brain can be in, in the way that liquidity and solidity are states that water can be in.

It is his backing of levelism —conscious states are caused by the lower level neuronal process in the brain— that has made Ned Block call Searle's Biological Naturalism «Non-reductive Physicalism» or «Default Physicalism» [Block (2002), p. 398]. Non-reductive physicalism has been the most influential and widely shared view about the relationship between «higher-level» properties and their underlying «lower level» properties. Ned Block agrees with Searle that the fact that physically different realizations of human functional organizations are conscious is not an *a priori* matter and could be said to depend on whether their brains have equivalent causal powers to ours in the sense of having the power to be physical basis of conscious states.

Again and again, Searle argues that conscious states exist insofar as someone experiences them and they have «first-person ontology» and that they are distinct from physical phenomena that have «third-person ontology». Nonetheless, even if we can trace some non-reductionistic and dualistic features in Biological Naturalism, Searle is neither a complete non-reductionist nor a complete dualist. For, non-reductionism often takes on dualism, which is still an option for Searle, although levelism rejects dualism without disposing reductionism as a live option.

The option of reductionism is as much open to him as the option of non-reductionism. Like the reductionist, Searle argues that mental states are caused by physical states. However, he cannot exercise the option of reductionism, because consciousness cannot be causally reducible to the brain process. That is to say, a perfect science of the brain would still not lead to an ontological reduction of consciousness in the way that our present science can reduce heat, solidity, colour or sound. It is also that physicalism, in both its reductive and non-reductive forms, fails to deal adequately with the problems of properly characterizing the nature of consciousness itself, making intelligible the relation between consciousness and the «physical», and creating the intellectual space for a shift in philosophical framework that would enable us to deal adequately with the first two problems, either the first or the second. The diagnosis of this failure is connected to the fact that consciousness cannot be treated in its own terms while being simultaneously fitted into an object-based conceptual schema. This grants the leeway for appending Searle's theory of mental causation in the category of Soft-reductionism. Soft-Line reductionists are much more amenable to first-person questions. The soft-line reductionists believe that the *cause* of everything which is going on in the mind is the physical system, and this physical system is explainable from the third-person point of view, but there may still be some emergent phenomena which are not captured by a purely physical, third-person description [Chalmers (1990)]. In the same way, Searle provides a biological reason for a psychological explanation. For him, the existence of consciousness can be explained by causal interactions between elements of the brain at the micro-level. Consciousness is a higher level or emergent property in the sense in which solidity is a higher emergent property of H₂O molecules when they are in a lattice structure of ice, and liquidity is similarly a higher-level emergent property of H₂O molecules, when they are rolling around on each other as in water [Searle (1983), p. 89]. But consciousness cannot itself be deduced or calculated from the sheer physical structure of the neurons without some additional account of the causal relations between them. This casts light on the fact that Searle's Biological Naturalism espouses that there exists an explanatory gap at the physiological level.

III. MODIFIED EPIPHENOMENALISM AS A BETTER CHOICE

While supporting bottom-up causation with supervenience, Searle also asserts that once we recognize the existence of bottom-up, micro-to-macro forms of causation, the notion of supervenience no longer does any work in philosophy. Similarly, for him, a difference in macro states would not necessarily involve a corresponding difference in micro states. It seems that with soft-reductionism, Searle defends a modified form of epiphenomenalism, even though he himself rejects it. Epiphenomenalism holds that conscious experiences are events occurring in an immaterial substance, causally determined by events in the brain but having no physical effects. Although the links, which exist between brain events and conscious occurrences, are causal links, they cannot be causal links, which are subsumable under purely physical laws, says Searle. With these types of argument he opens a door for mysterianism also [Churchlands (1996), p. 211]. For the methodology of science, the chief significance of his layered model lies in the relationship that is thought to hold between properties at adjacent levels. The weakness of Searle's theory of mental causation lies in the fact that, he fails to establish such a relationship. Unless there must be some identifiable features between the lower and higher-order features, they cannot succeed to explain the exact mode of relationship. Later, Searle has claimed that the psychological processes, though they are themselves caused by lower level neuronal process, nonetheless do not provide sufficient causal conditions for the subsequent psychological event of intentional action. So there is a gap at the psychological level, but not in the form of bottom-up causation between the neurobiological level and the psychological level, and not at the neurobiological level between any state of the system and the next state of the system. This would create physiological determinism along psychological libertarianism. In this sense his theory is modified epiphenomenalism. Modified epiphenomenalism advocates that the psychological processes of rational decision making do not really matter. This implies that Searle's theory of mental causation invites an explanatory gap not only at the physiological level, but at the psychological level also. This becomes clearly evident in his later discussion on free will.

Searle's mental causation gets completed with the two important hypotheses, which he formulates in the context of the problem of freedom of the will. He accepts that this is essentially a problem about a certain aspect of consciousness, namely that form of consciousness that manifests these sorts of gap. The two hypotheses are stated as follows: (1) Psychological indeterminism coexists with neurobiological determinism; and (2) Psychological indeterminism is matched by neurobiological indeterminism. For Searle, the problem of freedom of the will arises for those parts of the conscious field in which we experience the gap. These are the cases traditionally called «volition». On this account, the problem of free will only arises for volitional or

active consciousness. The argument is pursued further in his book *Rationality in Action*, where Searle argues that the very operation of rationality presupposes the gap [Searle (2001), p. 261].

Even while accepting free will or voluntary action, which he identifies with volitional consciousness, he sets out to remark that the acceptance is not possible without gaps. He makes a distinction between seeing my left hand, which is a passive act and raising my left hand which has causal antecedents of free will. He admits a first gap between the reason for the decision and the making of the decision, and the second gap between the making decision and the actual onset of the action. Free will, as it is traditionally conceived, cannot be identified with anything else, according to Searle. In fact, he is much more specific in mentioning three gaps in the structure of normal, voluntary human action. The gap between the reflection on the reasons and the decisions constitutes the prior intention of action. The gap between prior intention and the actual initiation of action is the second, and the third is the gap in the execution of action through time. So the psychological antecedents of a human action cannot be said to be causally sufficient for the performance of the action. They function causally, but they do not function in virtue of being causally sufficient. They are causally antecedent but not causally sufficient. Even if Searle rebuffs modified epiphenomenalism, this is a clear example of it. That is to say, the above arguments imply that mental states cannot have any physical effects, the fundamental assumption on which the epiphenomenalism rests.

Concisely, we can have an equally plausible positive and negative attitude towards both views: explanatory gap is either bridgeable or unbridgeable. On the positive side, the explanatory gap could be closed so long as a mental theory of causation is made to work. But, Biological Naturalism is less successful in exercising a complete theory of mental causation. On the negative side, what falsifies the latter however is the way the intentional theory of mental causation with its entire widened base converges on a view of rationality and a realistic view of social reality, which Searle accepts later in his book *The Construction of Social Reality*. In this book, he defends external realism that shows there is a reality that exists totally independent of our representations of it. And, this is a clear indication of his embracing modified epiphenomenalism and realism together. All these highlight the existence of explanatory gaps, which he himself admits because the explanation of mental phenomena is inadequate. Still, we can hope that once the proper conceptual framework has been established, a purely physical explanation will eventually be achieved.

*Department of Humanities and Management
Birla Institute of Technology and Science Pilani
Goa Campus, Goa, India – 403 726
E-mail: reennaac@yahoo.com*

REFERENCES

- BELOT, G. & EARMAN, J. (1997), «Chaos out of Order: Quantum Mechanics, the Correspondence Principle and Chaos», in *Studies in History and Philosophy of Modern Physics*, 2, Amsterdam, Elsevier, pp. 147-182.
- BLOCK, N. (2002), «The Harder Problem of Consciousness», *The Journal of Philosophy*, Vol. XCIX (8), pp. 391-425.
- CHALMERS, D. (1996), *The Conscious Mind: In Search of a Fundamental Theory*, Oxford, Oxford University Press.
- CHURCHLAND, P. and CHURCHLAND, P. (1996), «Rediscovery of Light», *Journal of Philosophy*, vol. 93, pp. 211-228.
- CRICK, F. and KOCH, C. (1990), «Toward a Neurobiological Theory of Consciousness», in *Seminars in Neuroscience* 2, Academic Press, pp. 263-275.
- KIM, J. (1996), *Philosophy of Mind*, Boulder, CO, Westview Press.
- (2000), *Mind in a Physical World. An Essay on the Mind-Body Problem and Mental Causation*, Cambridge, Massachusetts, The MIT Press.
- LEVINE, J. (1983), «Materialism and qualia: the explanatory gap», in *Pacific Philosophical Quarterly*, 64, pp. 354-361.
- NAGEL, T. (1974), «What is it like to be a Bat?», in *Philosophical Review*, 83, pp. 435-456.
- POPPER, K. (1994), *Knowledge and the Mind Body Problem: In defense of Interaction*, M. A. Notturmo, ed., London, Routledge.
- RAMACHANDRAN, V.S. (2006), «In the Mind of the Brain. A conversation by Sashi Kumar», in *Frontline*, an Indian National Magazine, April 7, pp. 5-6.
- SEARLE, J. (1984), *Minds, Brains and Science*, Cambridge, Mass., Harvard University Press.
- (1992), *The Rediscovery of the Mind*, Cambridge, Mass., The MIT Press.
- (2000), «Consciousness, Free Action and the Brain», in *Journal of Consciousness Studies*, 7, nr. 10, pp. 19-48.
- (2001), *Rationality in Action*, Cambridge, Mass., The MIT Press.
- STURGEON, S. (1998), «Physicalism and Overdetermination», *Mind*, 107, pp. 411-432.