

# Against McGinn's Mysterianism

Erhan Demirciođlu

**Abstract:** There are two claims that are central to McGinn's *mysterianism*: (1) there is a naturalist and constructive solution of the mind-body problem, and (2) we human beings are incapable in principle of solving the mind-body problem. I believe (1) and (2) are compatible: the truth of one does not entail the falsity of the other. However, I will argue that the reasons McGinn presents for thinking that (2) is true are incompatible with the truth of (1), at least on a fairly standard conception of the terms 'naturalist' and 'constructive', which McGinn himself seems to take for granted.

**Keywords:** Mysterianism, the mind-body problem, cognitive closure, Colin McGinn

**Özet:** McGinn'in *gizemcilik* adı verilen görüşü açısından iki iddia merkezi önemdedir: (1) zihin-beden probleminin doğalcı ve yapıcı bir çözümü vardır ve (2) insanlar zihin-beden problemini ilkesel olarak çözemezler. (1) ve (2), çelişik iki tez değildir: birinin doğruluđu diğerinin yanlış olmasını gerektirmez. Fakat savunacağım iddia odur ki, McGinn'in (2)'nin doğruluđuna dair verdiđi gerekçeler (1)'in doğruluđu ile – McGinn'in kendisinin de varsaydıđı, 'dođalcı' ve 'yapıcı' terimlerinin standart yorumlarını hesaba kattığımızda – çelişik durumdadır.

**Anahtar Kelimeler:** Gizemcilik, zihin-beden problemi, bilişsel kapalılık, Colin McGinn

There are two claims that are central to McGinn's *mysterianism*: (1) there is a naturalist and constructive solution of the mind-body problem, and (2) we human beings are in principle incapable of solving the mind-body problem. The combination of these two claims seems hard to digest. The falsity of (1) would be a plausible explanation of (2): it would be no wonder that we cannot provide a naturalistic solution to a problem for which no naturalistic solution exists. However, if (1) is true, how are we to account for the truth of (2)? Or conversely, does not the truth of (2) provide us with a good reason to believe that (1) is false?

I believe that (1) and (2) are compatible: the truth of one does not entail the falsity of the other. However, I will argue that the reasons McGinn presents for thinking that (2) is true are incompatible with the truth of

(1), at least on a fairly standard conception of the terms 'naturalist' and 'constructive', which McGinn himself seems to take granted.

McGinn defines the mind-body problem that he claims we cannot solve as follows:

How is it possible for conscious states to depend upon brain states? How can technicolor phenomenology arise from soggy grey matter? What makes the bodily organ we call the brain so radically different from other bodily organs, say the kidneys – the body parts without a trace of consciousness? How could the aggregation of millions of individually insentient neurons generate subjective awareness? We know that brain are the de facto causal basis of consciousness, but we have, it seems, no understanding whatever of how this can be so (pp. 394-5).<sup>1</sup>

The problem here concerns what is commonly called *phenomenal* consciousness, which McGinn himself calls "the hard nut of the mind-body problem" (p. 394). It is the problem of understanding how the (phenomenally) conscious mind is related to the brain. McGinn grants that we know enough about the mind-body relation to know that conscious mental states are caused by brain states but the question is how we are to explain this very relation that we know holds between those different states. The relation between the conscious mind and the brain "strikes us as miraculous, eerie, even faintly comic" (p. 395) and the mind-body problem is "the problem of understanding how the miracle is wrought, thus removing the sense of deep mystery" (p. 395).

McGinn's initial diagnosis of our bafflement about the possibility of there being a causal relation between consciousness and the brain is that "[n]eural transmissions just seem like the wrong kind of materials with which to bring consciousness into the world" (p. 395). The astonishment we feel about the possibility of a causal relation between consciousness and the brain is often expressed by a question of the form "how can *this* (consciousness) be caused *that* (the brain)?" The problem is that brains do not seem to be radically different from other bodily organs in a way that would support an explanation of how the All page references are to McGinn (1989), unless otherwise noted. former but not the latter give

---

<sup>1</sup> All page references are to McGinn (1989), unless otherwise noted.

rise to conscious mental states. Of course, a brain is different from, say, a kidney in term of its physical features and structures; however, the point is that *the way it is different* seems not to provide an explanation of what makes it responsible from consciousness.<sup>2</sup>

According to McGinn, the sense of mystery can be removed by having a theory that “describes the link between consciousness and the brain in a way that is no more remarkable (or alarming) than the way we now describe the link between the liver and bile” (p. 362). That the liver secretes bile might have been thought of as magical or mysterious to some of us who does not have a theory that explains how this is possible. Thanks to our current level of anatomical knowledge, there is now nothing magical for us about the secreting relation that holds between the liver and bile. McGinn holds that a theory that is intended to resolve the sense of mystery attached to the causal relation between consciousness and the brain needs to meet a similar adequacy constraint: it must explain how the brain “secretes” consciousness in a way relevantly similar to the way an anatomical theory explains how the liver secretes bile.<sup>3</sup>

McGinn argues that our situation with respect to an explanatory theory of the relation between consciousness and the brain is such that the theory is in principle beyond our cognitive reach. “What I want to suggest is that,” McGinn writes, “the nature of the psychophysical connection has a full and non-mysterious explanation in a certain science, but this science is inaccessible to us a matter of principle” (p. 401). A thesis that is central to McGinn’s overall position is that we are precluded in principle from understanding the correct scientific theory that accounts for the causal nexus between consciousness and the brain. Let us call this thesis *the inaccessibility thesis*.<sup>4</sup>

McGinn articulates the inaccessibility thesis in terms of the notion of cognitive closure, which he defines as follows:

A type of mind *M* is cognitively closed with respect to a property *P* (or theory *T*) if and only if the concept-forming procedures at *M*’s disposal cannot extend to a grasp of *P* (or an understanding of *T*) (p. 395).

<sup>2</sup> For further discussion, see Demircioğlu (2015).

<sup>3</sup> Note the famous materialist dictum of the 18th century French physiologist Pierre Cannabis: “The brain secretes thought as the liver secretes bile.”

<sup>4</sup> The inaccessibility thesis entails (2) on the plausible assumption that solving the mind-body problem requires understanding the relevant correct scientific theory. For the purposes of this paper, I will take (2) and the inaccessibility thesis to be equivalent.

According to McGinn, “we are cognitively closed with respect to that property [*P* of the brain that accounts naturalistically for consciousness]” (p. 396). On his view, we are cut off from achieving a conception of that natural property of the brain that accounts for the psychophysical link. We cannot, in McGinn’s own words, “grasp,” “represent,” “specify,” “identify,” “understand” or “arrive at” that natural property. Further, McGinn holds that our cognitive closure with respect to *P* entails and explains our cognitive closure with respect to *T*: it is because we cannot access to *P* that accounts for the psychophysical link that we cannot access to *T* that makes an essential appeal to *P*.<sup>5</sup>

Can there really be some properties (or concepts of those properties) to which we are cognitively closed? If the answer is no, then the inaccessibility thesis is false. However, the answer seems to be a clear yes if we reject a sort of idealism according to which to be is to be conceived. If the world exists independently of our conceptual capacities, then it is surely *possible* that some of its features are beyond our cognitive reach. I believe McGinn is completely within his rights when he makes the following claims about the possibility of cognitive closure:

Only a misplaced idealism about the natural world could warrant the dogmatic claim that everything is knowable by the human species at this stage of its evolutionary development (consider the same claim made on behalf of the intellect of cro-Magnon man) (p. 397).

Nothing, at least, in the concept of reality shows that everything real is open to the human concept-forming faculty – if, that is, we are realists about reality (p. 395).

McGinn is right to think that realism entails that there might be some properties that are cognitively inaccessible by us. If realism is correct, then there is no guarantee that the limits of our minds coincide with the limits of reality; and, if they do not coincide, then we are cognitively closed with respect to some properties instantiated by reality. It appears that the possibility of cognitive closure simply follows from realism and some innocuous assumptions about a particular object’s being independent

---

<sup>5</sup> McGinn writes: “That *P* is (as we might say) noumenal for *M* does not show that *P* does not occur in some naturalistic scientific theory *T* – it shows only that *T* is not cognitively accessible to *M*” (p. 395).

from some other object's conceiving it.

If, assuming "realism about reality," cognitive closure is possible, then the psychophysical link appears to be at least as a good candidate for being a feature to which we are cognitively closed as any other feature of the world:

It *may* be that every property for which we can form a concept is such that *it* could never solve the mind-body problem. We *could* be like five-year old children trying to understand Relativity Theory (p. 397).

McGinn's paper involves a rich stock of colorful possibility claims: we *might* be like traditional theologians finding themselves conceding cognitive closure with respect to certain of the properties of God, *or* like a Humean mind working with strictly empirical principles and yet trying to understand the physical world, *or* like a creature without spatial concepts trying to understand the possibility of motion.

Realism (about reality) entails the possibility of cognitive closure with respect to the conscious mind, and the possibility of that sort of cognitive closure means that (1) and (2) are compatible: the mind-body problem might have a naturalistic and constructive solution to which we have no cognitive access as a matter of principle. The relation between the conscious mind and the brain might be accountable for in naturalistic terms by a scientific theory without that theory being a possible object of our cognition.

Of course, to say that it is *possible* that our minds are cognitively closed to us is not to say that our minds *are* cognitively closed to us. Obviously, possibility does not entail actuality. Realism may move some significant way towards establishing the former but we need some extra reasons to think that what is claimed to be possible is actual. Realism tells us that some of the features of the world *might* be cognitively closed to us; but it does not tell *whether* there are actually any such features and, if there are, *which* features they are.

What reasons does McGinn provide us to hold that the inaccessibility thesis is true (and not merely possibly true)? McGinn's argument develops at three stages. First, he argues that we cannot get to *P* through introspection or by analyzing concepts that are introspection-based. Second, he argues

that we cannot get to *P* through a systematic empirical study of the brain or by making inferences from what that empirical study might possibly provide us with. Finally, given that these are the “two possible avenues open to us in our aspiration to identify *P*” (p. 397), we cannot get to *P* full stop.

I will not be concerned with the first and third stages of the argument (though this is not to say that interesting issues do not arise at those stages). I will focus on the second stage and argue that the reasons McGinn provides us with for the futility of empirical research exclude the possibility of a naturalistic solution of the mind-body problem.

At the second stage, McGinn argues *inter alia* that we are perceptually closed with respect to *P*: *P* is an imperceptible feature of the brain (or “is noumenal with respect to perception of the brain” (p. 398)). We cannot perceive *P* by “looking harder at” (p. 399) the brain.

It is clear that the imperceptibility of *P*(partially) explains why, as McGinn claims, we cannot get to *P* through empirical studies of the brain: the obvious connection is that empirical studies (partially) depend upon or are restricted by what is perceivable by us. The important question is, however, why we should believe that *P* is imperceptible: how on earth can there be a property of the brain that is principally closed to perception?

McGinn’s argument for the imperceptibility of *P* rests on the following conditional premises as well as the premise that *P* explains the psycho-physical link:

(P-S) If *P* is perceptual, then *P* is spatial.

(S-NE) If *P* is spatial, then it cannot explain the psycho-physical link.

In support of (P-S), McGinn writes:

[T]he senses are geared to representing a spatial world; they essentially represent things in space with spatially defined properties... The senses are responsive to certain *kinds* of properties – those that are essentially bound up with space... Kant was right, the form of outer sensibility is spatial (p. 399).

In support of (S-NE), McGinn writes:

[I]t is precisely *such*[spatial] properties that seem inherently incapable of resolving the mind-body problem: we cannot link consciousness to

the brain in virtue of spatial properties of the brain. There the brain is, an object of perception, laid out in space, containing spatially distributed processes; but consciousness defies explanation in such terms. Consciousness does not seem made up out of smaller spatial processes; yet perception of the brain seems limited to revealing such processes... [N]o spatial property will ever deliver a satisfying answer to the mind-body problem (p. 399).

The point I want to emphasize is that on McGinn's account, *P* is not only imperceptible but also non-spatial. The premise that *P* explains the psycho-physical link and (S-NE) together entail the non-spatiality of *P*.

In the final analysis, the overall picture that emerges from McGinn's argument for the inaccessibility thesis is as follows. The inaccessibility thesis is true, according to McGinn, on account of the fact that the feature *P* of the brain that explains the psychophysical link is not open to (human) perception. The reason why *P* is perceptually closed to perception is that no conceivable perceptible feature can explain the psychophysical link, and this is in turn because no conceivable spatial feature can explain the psychophysical link. The move from perceptibility to spatiality is based on the idea that all perceptual features are spatial. So, a basic reason McGinn provides for the truth of the inaccessibility thesis is that *P*, that feature of the brain that explains the psychophysical link, is non-spatial.

I believe there are good reasons to think that the non-spatiality of *P* provides some (at least *prima facie*) support for the inaccessibility thesis. It is highly plausible that if *P* is non-spatial, then it is not perceivable (imagining perceiving non-spatial things is very hard, to say the least). Further, if *P* is not perceivable, then what might initially seem to be a serious candidate for being the route to *P* turns out not to be a candidate at all.

However, the problem I want to press is that if what makes the inaccessibility thesis (or (2) above) true is, at least in part, that *P* is non-spatial, then it is hard to see how (1) can be true, i.e. how there can be a naturalist and constructive solution of the mind-body problem.

How can there be a *naturalist* solution of the mind-body problem if the property that explains the link between the two is not spatial? In order to appreciate the force of the question, one only needs to pay some

attention to the reason why we take, for instance, Cartesian dualism to be a non-naturalistic theory. Consider, for instance, McGinn's own remarks about "supernatural" solutions:

[One] form [of the solutions to the mind-body problem], which has been historically dominant, frankly admits that nothing merely natural could do the job, and suggests instead that we invoke supernatural entities or divine interventions. Thus we have Cartesian dualism and Leibnizian pre-established harmony (p. 395).

It is clear that Cartesian dualism holds that nothing merely natural could do the job if this means that nothing spatial could do the job. Cartesian dualism maintains that there is no naturalist solution of the mind-body problem if a naturalist solution requires that mental states (or thoughts) be the states of a spatial object. The root idea of Cartesian dualism is that the solution of the mind-body problem is possible only on the condition that mental states are taken to be the states of a *non-spatial* stuff.

So, the reason why we do not classify Cartesian dualism as a naturalist solution is that it is committed to the thesis that nothing spatial could do the job of solving the mind-body problem. Now, if this is so, then by parity of reasons, McGinn's mysterianism that holds that nothing spatial could do the job is also committed to the thesis that there can be no naturalist solution of the mind-body problem. The upshot is that McGinn's invoking the idea of the non-spatiality of *P* undermines his initial commitment to a naturalist solution of the problem.<sup>6</sup>

How can there be a *constructive* solution of the mind-body problem if the property that explains the link between the two is not spatial? To see how McGinn himself conceives the idea of a constructive solution, consider the follow passage:

One form [of the solutions to the mind-body problem], which we may call constructive, attempts to specify some natural property of the brain (or body) which explains how consciousness can

<sup>6</sup> Of course, Cartesian dualism is a sort of substance dualism but not merely a sort of property dualism. However, what makes it "non-naturalist" is not simply the fact that it is a form of substance dualism but the non-spatiality of one of those two putatively fundamental substances. And if this is true, then a sort of property dualism that advocates the non-spatiality of one of the fundamental properties must also count as "non-naturalist." I would like to thank Eylem Özalton for pressing on this issue.



be elicited from it. Thus functionalism, for example, suggests a property – namely, causal role – which is held to be satisfied by both brain states and mental states; this property is supposed to explain how conscious states can come from brain states (p. 395).

McGinn holds that a constructive solution proceeds by identifying a property that is shared by both kinds of states. I believe this is reasonable: there can be no constructive account of how a state can give rise to another, entirely heterogeneous state. A necessary condition for a constructive solution appears to be there being at least one common property shared by different kinds of states.

Now, if *P* is non-spatial, then it is not clear that there is any feature that it shares with spatial properties of the brain and that can be deployed in a constructive account.<sup>7</sup> Spatiality appears to be such a fundamental kind to which properties might belong that no property that does not fall under that kind can be given a constructive account in terms of those properties that do fall under it. Being spatial appears to be such a generic condition (as McGinn himself alludes to, something like a Kantian form) that no two properties can share some interesting features unless they are both spatial. The idea that *P* is non-spatial cuts the ties it might have had with the ordinary, spatial properties of the brain so severely that whatever hopes one might have had for a constructive solution appears to be dashed.

The point to be emphasized is that, on McGinn's account, *P* is not merely conceived as non-spatial but is non-spatial. McGinn does not argue that there is a spatial property of the brain that we somehow inescapably conceive as non-spatial due to the peculiarity of our cognitive mechanisms but he argues that *P* is non-spatial. McGinn makes a point about the *metaphysical* nature of *P*, not merely a point about our epistemic situation with respect to *P*. In fact, on his account, what explains our cognitive closure with respect to *P* is the very metaphysical nature of *P*, its non-spatiality.

If McGinn's point about *P* concerns its metaphysical nature, that is, if *P* is non-spatial and not merely a property that we conceive as non-spatial, then I believe the remarks above pointing at the impossibility of a

---

<sup>7</sup> Of course, there are some uninteresting features like *being a property* that is had both by *P* and spatial properties of the brain. Evidently, however, they are unhelpful when it comes to providing a constructive solution to the mind-body problem.

constructive explanation of  $P$  in terms of the spatial properties of the brain stand to reason. The possibility of a constructive explanation would not be excluded if  $P$  were a spatial property that is merely conceived as non-spatial but, as I have argued, this is not the case on McGinn's view.

The upshot is that the thesis of the non-spatiality of  $P$ , upon which McGinn's defense of (2)(or the inaccessibility thesis) rests, entails the falsity of (1), another central thesis of McGinn's mysterianism: if  $P$  is non-spatial, then there is no naturalist and constructive solution of mind-body problem. I agree with McGinn that (1) and (2) are consistent but this is not to say that every thesis that one might appeal to for defending (2) accords with (1). The idea behind the thesis of the non-spatiality of  $P$  is that nothing spatial can generate consciousness, and this very idea is what motivates the Cartesian dualism to stipulate the existence of non-spatial stuff. If the reason why we do not call the Cartesian dualism as naturalistic is its commitment to the thesis that nothing spatial can do the job, as it seems to be, then no account like McGinn's that makes the same commitment can reasonably hope for a naturalistic solution. Further, there can be no constructive solution of the mind-body problem if  $P$  is non-spatial. This is because if  $P$  is non-spatial, then  $P$  can share no (interesting) features with the spatial features of the brain and also because if there are no common features, then there can be no constructive solution.

---

An earlier version of this paper was presented at a Koç University Philosophy Department colloquium in December, 2015. I would like to thank all the participants for their helpful suggestions. I am grateful also to an anonymous referee for valuable comments.

---

Erhan Demircioğlu, Koç Üniversitesi, Türkiye

### **References**

- Demircioğlu, E. (2015). The puzzle of consciousness. *Cilicia Journal of Philosophy*, 2, 76-85.
- McGinn, C. (1989). Can we solve the mind-body problem? *Mind*, 98 (391), 349-366.