Mcginn’s Theory Of Consciousness, 
and Searle’s Indignant Response

Darren Domsky
York University, Toronto

ABSTRACT

Colin McGinn, in The Problem of Consciousness, defends the radical view that human beings in principle cannot solve the mystery of consciousness. He claims that, because consciousness has a structure that is cognitively sealed from human beings, we are incapable of understanding the riddle of the conscious mind. John Searle, in The Rediscovery of the Mind, rejects McGinn’s understanding of mind, and offers a radically different account. Consciousness, he explains, is simply a system feature of the brain, and is neither hidden nor incomprehensible. I contrast and assess the two positions, and conclude that McGinn’s stance is simply untenable.
In his *The Problem of Consciousness*, Colin McGinn puts forward an interesting
theory of how mind and body connect. He claims that an unknowable property P, a
property of consciousness, somehow is responsible for the mind-body connection. In
his opinion, consciousness itself has a hidden structure, one which will always be in
principle unknowable by human beings. Citing logical requirement, intuitive assurance,
theoretical necessity, and agreement with empirical findings, McGinn argues that we
have no choice but to agree with his theory. John Searle, it would seem, would beg to
differ. In *The Rediscovery of the Mind*, he opposes McGinn quite neatly, giving a
radically different picture of consciousness, as something which is knowable, open to
us, unhidden, and the feature of a material system which happens to have the causal
powers of the brain. The first purpose of this paper will be to flush out both
philosophers’ arguments for their respective positions, and to accurately understand
why it is they believe what they do about the mind. The second will be to analyse them
comparatively, to contrast their strengths and weaknesses, and to find which is the
stronger theory of consciousness.

McGinn’s first, and, in his opinion, most convincing reason for crediting
consciousness with a hidden structure is that it would explain the logical properties of
consciousness. In his opinion, the prominent logicians of a century ago set an
uncannily useful precedent. In order to account for how propositions expressed by
sentences in a natural language seem to require something transcendental to explain
how they present themselves to us, Frege, Russell, and the early Wittgenstein believed
that they had to postulate the existence of just such a hidden structure. Because the
grammar, the very surface, of logical language did not seem to account for the actual
logical relations, it seemed necessary to assume that something else existed beneath
the surface, linking the relations to the grammar. McGinn believes that a similar
surface/deep distinction is required in the philosophy of mind. In order to link mind to
matter, McGinn claims that we logically need to postulate the existence of a property P
in the world, a property “responsible for the capacity of matter to form the basis of
consciousness, or for the capacity of consciousness to take its rise from matter”
(McGinn, 58). McGinn believes we have good reason to think that this logically required
property is radically unknowable by us in principle. Because the two radically dissimilar
kinds of entity are clearly correlated with one another, but because we cannot
understand how the mind can “be physically governed ... and yet be so utterly unlike
that which governs it” (McGinn, 100), McGinn claims that the situation must remain for
us a mystery. He maintains that, by necessity, we will never be able to “explain how it is
that chunks of matter can develop an inner life” (McGinn, 104). The “logical properties
of conscious thoughts” (McGinn, 91) force us to believe that there is a connector, but
our inability to conceive of what it might be should convince us that it must be
permanently hidden. What is more, McGinn believes that the hidden structure must be
vitaly linked to consciousness itself, so much so that, in his opinion, the only sensible
possibility is that the hidden structure must belong to consciousness. He contends that
“conscious states possess a hidden natural (not logical) structure which mediates
between their surface properties and the physical facts on which they constitutively
depend” (McGinn, 100). For McGinn, logic does not merely give him an analogy of the
mind; it gives him a blue print. According to McGinn:

We are like thwarted logicians who know that the surface structure of
sentences is inadequate to explain their logical properties but who cannot
for the life of them develop the logical concepts needed to account for these properties: They know that a hidden logical structure must exist, since the surface of sentences will not do the job -- and there must be an answer to their puzzles ... (McGinn, 120)

The analytical philosophers of a century ago cut the logical mould, but McGinn’s property P, unknowable and yet responsible for the connection between brain and consciousness, fits it perfectly. Thus, McGinn believes that we are logically compelled to give consciousness a secret underbelly, a property P which accounts for the mind-body connection and which necessarily remains unknowable for humans.

Beyond its logical necessity, however, McGinn argues that we also have very good intuitive reason for believing that consciousness has a hidden structure, inaccessible to introspection. For one thing, it is not difficult to show how weak our powers of introspection can be. Since introspection, for McGinn, is the doorway to consciousness, this weakness is crippling for any theory that holds consciousness to be completely knowable by the subject. Showing that introspection is limited seems especially simple. In the first place, introspection offers us no information on the consciousness of others, and it certainly tells us nothing about the causal picture behind our different experiences. Also, introspection cannot even begin to form “a conception of alien forms of consciousness; and introspection is utterly incapable of informing us of the causal background of our conscious states.” (McGinn, 78) For McGinn, this latter point is extremely significant:

... there is a colossal amount of cognitive processing that goes on of which we are not aware. This processing is typically described in the language of computation and information-processing; it is not neat neural language. For example, the computational processes that underlie vision, rich and complex as they are, are hidden to introspection. This is why the process of seeing something strikes us as a lot simpler than it really is. (McGinn, 77)
If introspection seems impotent or has sharp limits, it seems likely, says McGinn, that consciousness, which is examined through introspection, might partially be the culprit. If introspection is limited, and McGinn believes that we have good reason to think that it is, it seems reasonable to suppose that consciousness itself has hidden components that are in principle unknowable by us. What is more, the idea of hidden structure, McGinn claims, is not alien to us. After all, it is not difficult for us to accept hidden structure in the matter that surrounds us:

Thus liquids, for example, have a hidden structure that explains their characteristic macro-properties; why they flow as they do, assume the shape of their container, turn solid in certain conditions. ... The macro properties of liquids emerge, as we might say, from this hidden structure. (McGinn, 79)

According to McGinn, it is no great intuitive adventure to suppose that the hidden variables commonly postulated in “physics, biology, linguistics, [and] psychology” (McGinn, 89) are real and effective in the world, even if they also seem radically unknowable. Indeed, McGinn observes that, in the ordinary world, we often “postulate the existence of a level of reality whose actual constitution we admittedly cannot specify.” (McGinn, 119) It seems that nature contains a vast series of hidden structures which make the world behave the way it does, and so postulating a hidden structure within consciousness itself, according to McGinn, seems intuitively natural. Thus, McGinn goes beyond stating that a hidden structure within consciousness is logically required; he appeals to our intuitions, and argues that the possibility is completely synchronised with them.

Of course, McGinn does not stop there, either. He argues not only that a hidden underside of consciousness is logically necessary and intuitively plausible, but also that
it performs “an indispensable theoretical role.” (McGinn, 65) McGinn is fairly prudent. He realises that if we “have no ready explanation to hand of what the hiddenness of consciousness might consist in or arise from, [then] we are apt to overlook the possibility that a proportion of it might be hidden” (McGinn, 65), and perhaps with good reason. However, he argues that a hidden structure of consciousness can fill the theoretical gap between mind and body, and hence do us the service of providing a solution to the mind-body problem. He attests that

... we know that what we introspect of consciousness ... is insufficient to explain how consciousness relates to the physical world; so we have reason to think that consciousness has properties that go beyond this mode of access to it. (McGinn, 120)

For McGinn, the surface of consciousness does not provide the answer to our mind-body questions. Simply put, the surface has not and seems incapable of providing us with an understanding of how it is that mental and physical can join. McGinn’s deep layer of consciousness, quite pragmatically, takes care of the mind-body problem, and so it seems there is good theoretical reason to uphold it.

What we should do is pause and ask ourselves whether the surface exhausts the reality; for if it does not, then the hidden part might well contain what is needed to keep consciousness glued down to the physical world, where it belongs. (McGinn, 108)

For McGinn, his theory at least has the ability to account for the mind-body connection and therefore to play a particularly difficult theoretical role, whereas all competing theories do not even have the potential to do so.

Perhaps McGinn’s most intriguing reason for us to take up his belief that there is an underside to consciousness is because our best science has compelling empirical evidence for consciousness’ having a partially hidden structure. McGinn claims that if
we had “cases in which the surface properties [of consciousness] are absent and yet the usual sorts of effect transpire ... then we could attribute those effects to properties other than the surface ones.” (McGinn, 109) Blindsight, in his opinion, is just that sort of case. When people claim to be unable to see and yet pass tests of vision exactly as easily as a sighted person, McGinn believes they are demonstrating how our deep level of consciousness, which is hidden from our introspective awareness, can manifest itself in our behaviour. According to McGinn, our best science cannot help but conclude that there apparently are two kinds of property inherent in consciousness: “surface properties, which are accessible to the subject introspectively; and deep properties, which are not so accessible.” (McGinn, 111) The very fact that we can observe in a human being seemingly conscious behaviour, even when it is “shorn of its phenomenal surface” (McGinn, 111), is for McGinn extremely persuasive empirical proof that consciousness does have a dark side, one that he believes cannot even in principle be uncovered. In his opinion, a hidden structure within consciousness is not only required logically, sound intuitively, and useful theoretically; it is also backed up with convincing empirical evidence.

John Searle, in his stance on consciousness, could hardly be outdone as a foil to McGinn. Every substantial claim that McGinn makes, even those that he takes to be most obvious, is denied by Searle. He disputes McGinn’s avowal that it is logically necessary that there be a hidden aspect within consciousness, and he certainly disagrees with the claim that a valuable theoretical role is filled by that hidden aspect. According to Searle, McGinn is fundamentally misguided about what the mind is. McGinn subscribes to a dualist understanding of mind, where consciousness is
understood to be “a kind of ‘stuff.’” (Searle, 104) McGinn often makes claims like “It would take a supernatural magician to extract consciousness from matter” (McGinn, 45) or “Consciousness has to be recreated every time a new sentient organism springs up” (McGinn, 46), claims which directly insinuate that consciousness is a sort of substance or thing. Searle argues that McGinn’s theoretical role is only created when the stuff of consciousness has to be linked with ordinary matter. There only seems to be a “gulf between matter and consciousness” (McGinn, 120) because matter and consciousness are mistakenly taken to be distinct and independently real entities. For Searle, the Cartesian claim that mind and body are two separate entities, only “with the added disadvantage that the hidden structure of consciousness is unknowable in principle” (Searle, 104), is severely outdated and nonsensical. In fact, Searle believes that McGinn’s theory might even be less plausible than Descartes’, since property P is even less accessible than the pineal gland. First of all, postulating a hidden structure, according to Searle, “gets us nowhere”. (Searle, 105) After all, if “you need a link between consciousness and the brain, then you need a link between the hidden structure of consciousness and the brain” (Searle, 105), and so the hidden structure solves nothing. Not only does the postulation not solve anything, but, according to Searle, it is also erroneous. For Searle, unlike for McGinn, consciousness is a system feature of our brain, just as roundness is a system feature of objects and wetness is a system feature of water. Because of this, there really “is no ‘link’ between consciousness and the brain, any more than there is a link between the liquidity of water and H₂O molecules.” (Searle, 105) The logical necessity and the performance of the
theoretical role both evaporate when the mind and body are no longer viewed as distinct entities needing to be linked.

In addition to this, however, Searle believes that the application of the word ‘hidden’ to the word ‘consciousness’ is an error of its own. By definition, Searle explains, consciousness must be a subjectively experienced thing, something whose entire reality just is its surface appearance. According to Searle, consciousness has an irreducibly subjective quality, and so “we cannot make the appearance-reality distinction because the appearance is the reality.” (Searle, 122) The appearance of consciousness to us is all that consciousness is, even though it is created by the causal powers of the brain. Searle argues for the connection principle, the idea “that all unconscious intentional states are in principle accessible to consciousness” (Searle, 156), and it follows from this that if something is in principle inaccessible to consciousness, then it is surely not related to consciousness in any way. At best, such a thing would be merely a “nonconscious neurophysiological [state] of the brain capable of giving rise to conscious thoughts and to the sorts of behaviour appropriate for someone having those thoughts” (Searle, 167). It would not be a hidden part of consciousness, but rather a non-understood dispositional state of our neural network. If a mental state is not fully conscious, “then the only facts are the existence of neurophysiological states capable of giving rise to conscious thoughts and to the sort of behaviour appropriate for someone having those thoughts.” (Searle, 166) Clearly, McGinn’s very claim that there could be a hidden structure within consciousness flies in the face of everything Searle believes about the active and purely subjective system-feature mind.
Searle also disagrees with McGinn about what our intuitions are on the subject. With regard to the faculty of introspection, Searle all but accuses McGinn of philosophical trickery. McGinn claims that introspection is the tool we use to observe consciousness, just as our senses are used to perceive the outside world, and, since introspection is limited, McGinn concludes that we have reason to think consciousness has a hidden structure. Searle makes himself quite clear on the subject when he states that the “doctrine of introspection is a good example of what Wittgenstein calls the bewitchment of our intelligence by means of language.” (Searle, 105) According to Searle, introspection is nothing like perception because perception involves the distinction between subject and an object, whereas introspection does not.

The model of ‘specting intro,’ that is, the model of an inner inspection, requires a distinction between the act of inspecting and the object inspected, and we cannot make any such distinction for consciousness. (Searle, 105)

The capacity to introspect, modelled on vision, is simply misguided, and should not, according to Searle, be taken literally but only metaphorically. In response to McGinn’s claims that certain causes of conscious experience, such as the biomechanics underlying vision, are hidden from us, Searle asks why we should expect anything different, and, more importantly, why we should care. The causes of consciousness, namely the causal powers of our brains, need not be consciously experienced themselves if consciousness is to result. In his words:

We are not aware in conscious experience of either the spatial location or the dimensions of our conscious experience, but why should we be? It is an extremely tricky neurophysiological question, one we are a long way from solving, to figure out exactly what the locus of conscious experience in our brains is. (Searle, 105)
Searle’s point is that neurons can cause consciousness without that consciousness having to know what the neurons are doing. If there is a hidden side to something, then surely our intuitions are that it is neurophysiology that is hidden, and not consciousness itself. Clearly, Searle will not agree with McGinn that we have any intuitive reason to think that consciousness might have some hidden structure. Rather, Searle would argue that intuition is completely on his side.

In response to McGinn’s claim that the available empirical evidence supports his theory that consciousness has a hidden structure, Searle again flatly disagrees. The key to Searle’s understanding of consciousness is that, in principle, nothing can be hidden from it and still be a kind of consciousness. His attack against the unconscious, then, becomes even more fierce when the target becomes a hidden structure within consciousness itself. As a weapon against a deep unconscious, Searle rejects blindsight type examples, simply because for him “cases where there is a blockage of some sort” (Searle, 163) do not count as instances of an unconsciousness inaccessible in principle to consciousness. So long as “there is nothing ‘in principle’ inaccessible to consciousness about the phenomena in question, ... it is not a counterexample to the connection principle.” (Searle, 163) For Searle, blindsight is simply our neurophysiology performing certain mechanical tasks, without any connection with consciousness. Part of brain is performing the task of gather visual information, even though our consciousness is not involved, and that is all. According to Searle, McGinn’s claim that the available empirical evidence favours the hypothesis that consciousness has a hidden underbelly is not only false but ludicrous. The reason, Searle explains, is because the notions are simply contradictory. Empirically proving that consciousness
has an unknowable structure, for Searle, is like empirically proving that there are round squares. Even before you look at the evidence cited, Searle believes that you can dismiss it in advance. Thus, Searle again parts ways with McGinn, disagreeing with him completely on the subject of how our current empirical evidence supports the existence of a hidden structure of the mind.

How strong is Searle's attack? Certainly, the debate between he and McGinn is an involved one. As far as empirical evidence is concerned, I feel that the debate is at a standstill, with neither McGinn's nor Searle's theory being better supported. There simply seem to be many ways to encompass the blindsight findings. Certainly, McGinn may be correct in claiming that a hidden structure of consciousness is responsible for people apparently receiving visual stimuli without knowingly seeing those stimuli, but he may equally be wrong. Searle might be right to disregard cases of blindsight as cases of some sort of blockage, which have no effect on whether or not something is in principle inaccessible to consciousness, but again he might not. Both philosophers are convinced that the empirical evidence favours their own theory, and both seem to have a chance at being correct. There is simply no way to tell—the evidence, on this point, is too controversial. Until more research is done, it seems unlikely that either philosopher will be able to cite empirical findings to his advantage.

When the other arguments are examined, however, it would seem that Searle's position is generally more appealing. On the question of logical necessity, for instance, I hardly see the reason why a hidden structure is necessary to logically account for some surface phenomenon. Except in extremely academic cases, where logic itself is being dealt with, it seems that invoking hidden structures in order to explain features of
reality is a slippery slope, and a needless one at that. Certainly, if a hidden structure is required at all, it seems difficult to avoid requiring another hidden structure to account for the first one, and so on. On top of this potential for a vicious regress, it would seem that Wittgenstein is right when he says that “Explanations come to an end somewhere” (Wittgenstein, 1). It seems quite reasonable to believe that there are brute facts in the world. Searle’s claim that the mind has no explanation than that it is created by a brain, or something with a brain’s causal powers, seems very acceptable. In my opinion, Searle is right about McGinn’s argument from logical necessity; it simply does not follow.

With respect to introspection, I again cannot help but think that Searle’s attack on McGinn is well aimed. I do not believe that we can set up an appearance/reality distinction where the mind is involved, simply because consciousness, as Searle describes it, really is only its appearance. Introspection, as modelled after perception, does seem to be a mistaken way of describing our experience of our own consciousness, and so the flaws that McGinn attributes to introspection simply do not seem to affect consciousness. Of course, McGinn is right when he says that, introspectively, we cannot imagine how our minds arise from matter. Still, why should that point to a hidden structure in consciousness? More specifically, why should we interpret McGinn’s peculiar intuitions as anything more than the “guess” (Searle, 104) that Searle describes it as? What kind of introspective intuition could possibly lead to such a precise speculation amidst many possible options? If intuition favours anyone, it would seem to favour Searle. As a system feature, the mind seems at least feasible, but as a kind of stuff connected to the brain via the hidden structure of its own
consciousness, the mind goes for a field trip. If McGinn is truly describing intuitions, they would seem to be his alone.

Lastly, when it comes to filling in theoretical gaps, McGinn’s property P makes for rather fishy caulking. In the first place, Searle is right: postulating a hidden structure gets us nowhere. One might even be inclined to say that postulating an unknown property P to solve the mind-body problem is simply skirting the problem. Certainly it is no less puzzling than the original situation, and, worse yet, because it is touted as a solution, it has the potential to blind one to the fact that the original mystery remains a mystery. Second, it seems that McGinn does not really fill any theoretical gaps at all. According to McGinn, in recognising that property P exists and that it is cognitively sealed from us, he is filling the theoretical gap of how the mind and brain are connected. It would seem, however, that all he actually does is provide a label by which to better refer to the solution, leaving us not with closed theoretical gaps but instead with theoretical gaps that are easier to refer to. A similar move would be to let ‘x’ be the meaning of life, and then to claim to have made headway in the search for the meaning of life, since now, at least, you can refer to it. In the end, it seems more reasonable to deny, as Searle does, that there is any theoretical gap at all, and to say instead that the mind is simply a system feature of the brain. It seems more pragmatic to assert that the two things are not separate entities in the first place, and that they therefore do not need to be theoretically connected at all, than to accept McGinn’s theory on the basis that it successfully connects them. In reality, McGinn’s property P does not fill the theoretical gap, but instead only labels it, and, because of this, it seems unreasonable to accept his theory on the basis that it fills a theoretical role.
When all of the arguments are weighed, and we compare McGinn’s theory of mind with Searle’s, Searle certainly seems to come out on top. His solution to the mind-body problem has well thought out foundations and is extremely difficult to uproot. Still, although his position does seem weaker, McGinn would seem to have some important things to say to Searle, things that Searle would do well not to brush off. For one thing, if we are to believe that consciousness is a system feature of the brain, we should remember that it still has to be an absolutely amazing system feature. Whatever it is that makes it possible for consciousness to emerge from the brain is certainly remarkable, and, in talking about the mind as a system feature, it would be wise to remember this. McGinn might serve as a reminder that perhaps Searle’s solution, the way it is put, is too fantastic or over simplified to be true. After all, if the mind really is a system feature, exactly at what point does that feature emerge? Sorites’ paradox seems to threaten, since if one atom does not make a mind, and adding a single atom seems insufficient to ever suddenly create the causal powers that create a mind, then it would seem that a mind could never emerge as a system feature of a mere collection of atoms. Also, could McGinn’s property P not feasibly exist at a higher level? Might it be wise for Searle to consider that, if McGinn’s property P were thought of as a property of the universe, something which is responsible for there being the physical laws and materials needed for a mind to emerge in the world, perhaps we should subscribe to it? It seems that there might reasonably be a property R of the universe that causes the system feature ‘roundness’ to emerge from collections of atoms. Could there not be a property P that causes minds to result from a system of atoms? Certainly, McGinn’s theory would have to be severely modified, but perhaps there is something to what he
says about the nature of consciousness. As it currently stands, McGinn’s theory is certainly weaker than Searle’s, but Searle would do well to pay McGinn’s merits careful attention. Searle may win out, but he certainly does not tell the whole story.

It would thus seem that Searle’s theory of the mind is more appealing than McGinn’s. Certainly, Searle does not say everything that needs to be said about consciousness, but his ideas certainly seem more viable than McGinn’s. Still, neither fighter is down. Both philosophers have interesting things to say, and the dispute needs to be further resolved. As we watch the ongoing battle, however, it seems to me that Searle is currently in the best position for achieving eventual success, whatever that might consist of.
WORKS CITED:

