The Epistemic Approach to the Problem of Consciousness

Daniel Stoljar, ANU

Introduction

At the end of a *Time* magazine summary of the present state of thinking on the hard problem of consciousness, Steven Pinker writes:

> And then there is the theory put forward by philosopher Colin McGinn that our vertigo when pondering the Hard Problem is itself a quirk of our brains. The brain is a product of evolution, and just as animal brains have their limitations, we have ours…[and so we]…can't intuitively grasp why neural information processing observed from the outside should give rise to subjective experience on the inside. (2007, 6)

Concerning this theory, Pinker remarks (2007, 6): "This is where I place my bet, though I admit that the theory could be demolished when an unborn genius—a Darwin or Einstein of consciousness—comes up with a flabbergasting new idea that suddenly makes it all clear to us."

In fact, it is not so easy to interpret what Pinker is saying here. McGinn's view is that it is biologically impossible for humans to solve the hard problem (McGinn 1989; see also McGinn 1991, 2004).[1] A consequence of this is that there could be no Darwins or Einsteins of consciousness, not at any rate if that means a *human* Darwin or Einstein. And while McGinn does not rule out that a *non-human* (e.g. a Martian) Darwin or Einstein might solve the problem—indeed, it is consistent with his view that even the dullest Martian might do so—a further consequence of his view is that no Martian could make the solution clear to us humans, for that too is biologically impossible.

[1] McGinn uses the phrase 'mind-body problem' rather than 'hard problem' but we can assume that his view applies to both (if indeed there is a difference). I will describe in more detail what the hard problem is immediately below.

What then does Pinker think? When he says he would place a bet on McGinn's view, it sounds like he thinks it is probably true. But when he says that he admits McGinn's view *could*[2] be demolished, it sounds like he thinks it is probably false.

Despite this, there is no doubt that Pinker is highly sympathetic to the general direction of McGinn's position. Both suppose that we are ignorant at least for the time being of something important and relevant when it comes to the hard problem, and that this fact has a significant implication for its solution. They both hold, as I will put it, an *epistemic view* of the hard problem of consciousness.

Moreover, they are not alone in doing so. The epistemic view is a common one in the history of philosophical and scientific thought about the hard problem. One can detect it or something similar in Arnauld's (1641) criticism of Descartes, in Locke's (1689) suggestion that God may 'super-add' the capacity of thinking to matter, in Priestley's (1777) discussion of organized matter, in Du Bois-Reymond's (1872) distinction between ignoramus and ignorabimus, and in Russell's (1927) views on 'the causal skeleton of the world'. Versions of the view are also present in Chomsky's (1975) distinction between mysteries and problems, in Nagel's (1974) discussion of whether the pre-Socratics could understand the identity of matter and energy, and in Jackson's (1982) example of the super-slugs, all three of which are the immediate precursors of McGinn's paper.

I too am a proponent of the epistemic approach, though the kind of position I favour is different again from those advanced by these writers. In the first part of this paper, I will outline in elementary terms the version of the view I think most plausible. In the second part, I will respond to two objections to the view. The first says that, while we may be ignorant of various features of the world, we are not ignorant of any feature that is *relevant* to the hard problem. The second says that, even if the epistemic approach is true, properly understood it is not an *answer* to the hard problem; indeed, it is no contribution to that problem at all. In the concluding part, I will offer some brief reflections on why the epistemic approach, despite its attractiveness, remains a minority view in contemporary philosophy of mind.

Which Question is at Issue?

Suppose we are asking for the explanation of some natural phenomenon. It doesn't matter for present purposes what it is—let it be the origin of some disease, or the operation of a particular enzyme, or the fluctuation of wool prices in medieval England. Now suppose

[2] I take it 'could' is important here; if Pinker had said 'would' no interpretative difficulty would arise.

somebody—Bill, let's call him—announces that he has finally arrived at a theory. "Bravo and congratulations," we say, gathering round, waiting for further details. Imagine then our disappointment when Bill steps up to the podium and says, "My theory is…wait for it...I have none. I have no theory at all. In fact, I haven't a clue about what explains the phenomenon in question." Perhaps some of us with commend Bill for his honesty. But most of us will feel cheated. "It's fair enough to be ignorant", we might say, "but it's outrageous to call it a theory. Even more outrageous is that people get paid by universities to say such things."

Is the epistemic view any better than this? At first glance you may think not. Surely the hard problem of consciousness is just the problem of explaining what consciousness is. And surely the epistemic view is just the 'theory' that we don't know what that explanation is. If so, a proponent of that view is no better than Bill, and should be subject to the same kind of opprobrium.

But this impression is misleading. The hard problem is not a single problem but a cluster of related questions and issues. It is true that one question in the cluster—let us call it Question A—is, 'What is the explanation of consciousness?' But a different question—let us call it Question B—asks not *what* the explanation of consciousness is, but rather *whether* there is, or could be, a certain sort of explanation, a sort sometimes called a reductive explanation. The notion of a reductive explanation may be spelled out in different ways, but here we may assume that there is a reductive explanation of consciousness if and only if there is a set of physical facts—perhaps facts about the brains and bodies of sentient creatures—that necessitate the facts of consciousness.[3] So from this point of view, Question B is, 'Is there a physical fact, or set of such facts, that necessitate the facts about consciousness?'

From a logical point of view, Question A and Question B are distinct. Question B is a yes/no question, Question A is not. Question B explicitly concerns the notion of a reductive explanation, Question A does not. Nevertheless they are also related. For example, often but not always, when people raise Question A, they presuppose a positive answer to Question B; that is, they assume that there is some relevant set of physical facts, and are simply asking what these are. Moreover, if the answer to Question B is negative, this constrains the sort of answers we might give to Question A, since we could not in that case answer it by providing a reductive explanation.

---

[3] By 'fact' here I mean a true proposition; by 'necessitate' here I mean what is sometimes called metaphysical or logical necessitation; on this view one fact necessitates another just in case in all possible worlds in which the first obtains, the second does too. This notion of reductive explanation is similar to but is not quite the same as that discussed in Chalmers and Jackson 2001.

While the questions are distinct, one might think there is no need to discuss them separately. "Just work on Question A", you might say, "and the answer to Question B will come out in the wash." This advice is plausible on the surface but it forgets that there are persuasive reasons for supposing that there *must* be a negative answer to Question B, reasons which are apparently quite independent of Question A. The reasons I have in mind are the famous arguments in philosophy of mind about zombies, Mary, inverted spectra and the rest. If successful, these arguments all suggest the same conclusion, namely, that there is no reductive explanation and hence that Question B must have a negative answer.

The epistemic view, in the form I think is most plausible, is concerned with Question B, and in particular, with these arguments that it must be answered in the negative. In brief, its suggestion is that these arguments presuppose that we have a complete knowledge of the facts relevant to answering them—that "all the facts that concern us lie open before us" as Wittgenstein memorably put it in a related context (see Wittgenstein 1960, 6). Since, according to the epistemic view, that presupposition is false, the arguments fail, and we are free to suppose that there is a reductive explanation of consciousness.

I have said that the epistemic view is concerned with Question B. I haven't said that it is *not* concerned with A—a view may be concerned with two questions at once after all. Is it or isn't it? Well, suppose it is not; then the proponent of the epistemic view is clearly not in the position of Bill above, and that problem at least is answered. But you might reasonably insist that no genuine contribution to the hard problem can avoid dealing with Question A somewhere along the line. If so, it looks as if the proponent of the epistemic view is in the same position as Bill after all. This brings us close to the second of the two objections I mentioned at the outset, namely, the problem of whether the epistemic view is a genuine contribution to the hard problem at all. However, let me postpone this issue until we have looked in more detail at the view.

The Conceivability Argument

I spoke just now of the 'famous arguments in philosophy of mind about zombies, Mary, the inverted spectra.' Let's examine one of the central arguments in this group—namely, the conceivability argument—and how the epistemic view seeks to answer it. (Here my discussion will be very brief since there is a huge literature dealing with the ins and outs of these arguments.)[4]

[4] For a recent selection of this literature, see Alter and Howell 2012

The first premise of the argument is that it is conceivable or imaginable that there is a possible world identical to the actual world as regards all physical facts but is different from it as regards some fact about consciousness. The second premise is that if this is conceivable, it is possible. The conclusion is that this is indeed possible, and in consequence there can be no reductive explanation of consciousness.

The epistemic view has a two-part response to this argument. The first part supposes that there is a type of physical fact that is relevant to consciousness but of which we are ignorant; I refer to this part of the response as *the ignorance hypothesis* (see Stoljar 2006).[5] This part of the epistemic view is a contingent claim, since it makes a claim about our epistemic capacities and achievements, and the relation between these capacities and achievements and the world. But it is a contingent claim that is rather plausible—at least prima facie. Surely we can assume that there are facts about consciousness of which we are ignorant!

The second part argues that, *if* the ignorance hypothesis is true, the conceivability argument is unpersuasive. To illustrate this second part, let us focus on the quantifier phrase 'all physical facts' that occurs explicitly or implicitly in both premises of the conceivability argument. If the ignorance hypothesis is true, this quantifier may be understood in two ways: it might be interpreted so that the physical facts that are relevant but unknown to us are included in its domain; or it might be interpreted so that those facts are not included. Either way, the argument is unpersuasive. For suppose the relevant but unknown physical facts *are* included in its domain; in that case it is not true that we can conceive of the possibility in question, and the first premise of the argument is false. Suppose instead they are *not* included in the domain; in that case the conclusion of the argument shows only that the facts about consciousness come apart from *some* physical facts, not that they come apart from *all*. But by itself that conclusion is unsurprising; certainly it does not entail that consciousness is a fundamental element in nature, which is the conclusion that proponents of the conceivability argument are usually interpreted as wanting to establish.[6]

The Epistemic Approach and Other Responses

---

[5] It bears emphasis that the ignorance hypothesis concerns types of facts, rather than individual facts; for discussion of this point see Stoljar 2006, 70-72.

[6] For extensive discussion of this view, see Stoljar 2006 and references therein; see also Stoljar 2015.

How does this answer to the conceivability argument relate to other, more traditional, answers?

The dualist responds to the argument by accepting it, and then supposing that consciousness *is* a fundamental element of nature, rather like space, time and gravity (or at any rate like space, time and gravity are often taken to be). The proponent of the epistemic view is clearly not a dualist, since he or she doesn't accept the argument. Moreover, the dualist is naturally understood as inferring that consciousness is fundamental from the claim that it is not derivative on known features of matter. The proponent of the epistemic view denies this, pointing out that, from the premise that something is not derivative on known features of matter, it does not follow that it is not derivative on anything; hence it does not follow that it is fundamental.

The eliminativist responds to the argument by agreeing with the dualist, conditional on the existence of consciousness—and then denies the condition. According to the eliminativist, in other words, consciousness does not exist. The proponent of the epistemic view is clearly not an eliminativist, since nothing in that view denies that consciousness exists.

The materialist[7] responds to the argument by rejecting it, and asserting that materialism is true. Of course a proponent of the epistemic view rejects the argument as well—how then is this view to be distinguished from materialism?

Well, the *standard* materialist, the sort you usually find in the literature on these matters, denies the claim I mentioned above, the ignorance hypothesis. The standard materialist thinks that all the relevant facts are in, or at any rate that one can assume that they are in when discussing the conceivability argument. Rather than appealing to the ignorance hypothesis, the standard materialist tries to disarm the conceivability argument in a different way.

How does the standard materialist try to disarm the argument? There are a large number of ideas about this in the literature: the ability hypothesis, the phenomenal concept strategy, various acquaintance views, appeals to the necessary a posteriori, to the distinction between concepts and properties, to representationalism, and so on and so forth. I will not go into these ideas here. Suffice it to say that I am persuaded that none of them is successful; that is, none of them explains where the conceivability argument goes wrong (see, e.g. Stoljar 2006, 175-217). So at this point, I am in agreement with dualists and eliminativists who

[7] By 'materialist' I will mean 'non-eliminativist materialist'.

likewise think that none of these ideas can be made to work. Conclusion: standard materialism should be rejected. Further conclusion: if dualism and eliminativism are likewise rejected, the remaining option is materialism of a non-standard kind, a kind that combines materialism with the ignorance hypothesis.

Should a non-standard materialist be called a 'materialist' in the first place? That depends on an interpretative question that turns out to be surprisingly difficult, namely, what *exactly* you have to believe in order to count as a materialist. By very weak standards, it is sufficient to count as a materialist if you hold only that consciousness is not fundamental. By that standard a proponent of the epistemic view can certainly be called a materialist. By much stronger, but more historically accurate, standards, a materialist is someone who thinks that everything is determined by the physical, where the notion of the physical is understood in the light of a well-known historical theory—for example, the sort of atomism articulated by the Greeks and revived in the early (pre-Newtonian) renaissance period by philosophers such as Gassendi. By that standard, non-standard materialism is certainly not a kind of materialism—but then again nor are the versions of 'materialism' held by philosophers such as David Lewis or J.J.C Smart, who are usually thought of as the paradigm contemporary examples of materialist philosophers.

My own view is that observations like this cause a major problem for the attempt to provide a clear statement of materialism; that is why the interpretative question I mentioned is so difficult.[8] However, for present purposes, we can afford to be relaxed: if you adopt weak standards, non-standard materialism is indeed a version of materialism; if you adopt strong standards, non-standard materialism doesn't look much like materialism—but then again neither do the contemporary theses that usually go by that name.

Varieties of the View

We have set out the epistemic view, and explained its relation to other responses to the conceivability argument. Let me bring this brief sketch of it to a close by noting that there are various different versions of the view.

One version is that advanced by McGinn in the work Pinker comments on above. As we have seen, McGinn's view is that it is biologically impossible for us to solve the hard problem. As McGinn develops this idea, it contains some elements that distinguish it from

[8] For discussion of these issues see Stoljar 2010

the view we have been looking at. For one thing, he does not bring the issue of ignorance to bear on the conceivability argument and similar arguments quite in the way that I do; moreover, he thinks of these arguments as having a conceptual, rather than an epistemic, origin.[9] Still, if we transpose what he says into our framework, the result is something like this: (a) we are ignorant of some of the physical facts that necessitate the facts about consciousness; and (b) this ignorance is not something that can be overcome consistent with our biological nature. Following Owen Flanagan (1992), we may call this the 'mysterian' version of the epistemic view.

There is no doubt that mysterianism is interesting and well worth exploring. But there is no need for a proponent of the epistemic view to be a mysterian: one can assert (a) without asserting (b). After all, to say that there are unknown relevant properties of matter is not to say that these properties will remain forever unknown to humans or to rational agents in general—it is not to say that we are *cognitively closed* with respect to these properties, in McGinn's well-known phrase.

Indeed, when you focus on it this last suggestion is extremely speculative. You can say that we are cognitively closed with respect to something only if you also know two further things: what the limits of human knowledge are, and whether the thing in question falls outside those limits. But we know neither of those things, and that is why the epistemic view I favour leaves the issue of cognitive closure open.[10]

Another version of the epistemic view, which has been getting a lot of attention in recent years, is a view inspired by Bertrand Russell's (1927) *The Analysis of Matter*. In that book, Russell outlines an epistemological position according to which in empirical inquiry we are limited to knowing only the logical or causal structure of the world—structural facts, as we might put it in more contemporary language. But he also thought that it was possible that there are non-structural facts in addition to structural facts, and went on to suggest, in effect, that facts that necessitate consciousness might include these non-structural facts. Transposing this into our framework, the result is something like this: (a) we are ignorant of some of the physical facts that necessitate the facts about consciousness; and (b) the physical facts in question are non-structural facts that our epistemological framework does not permit us to know.

[9] For some discussion of these points, see Stoljar 2005.

[10] For an interesting exchange on whether we should even leave it open, see Kriegel 2003 and Demircioglu 2016.

Like mysterianism, the Russellian version of the epistemic view is interesting and well worth exploring. But again no proponent of the epistemic view needs to be a Russellian: one can assert (a) without asserting (b). For one thing, to endorse the ignorance hypothesis is not to presuppose any special epistemology like the one sketched by Russell—*any* epistemological theory at all must acknowledge possibilities of the kind the ignorance hypothesis makes salient. Moreover, Russell's epistemological position taken absolutely literally is quite implausible. Take fish, for example. On the face of it, we know plenty of facts about fish—e.g. that they are pleasingly symmetrical—that seem to be neither logical nor causal facts about them. Hence we know facts about fish that are not structural in the relevant sense. Of course, the phrase 'structural' is, like fish, extremely slippery. Someone might interpret it in such a way that facts about fish of the kind I just mentioned count. But there is trouble ahead for the Russellian version of the epistemic view if this line is adopted. For the broader you make the notion of 'structural', the narrower the notion of 'non-structural' becomes, i.e., since these are a package deal. Consequently, the less likely it is that these non-structural facts can play a role in necessitating the facts about consciousness.[11]

The Relevance Objection

So much for saying what the epistemic view *is*—turning now to objections, like any other position in philosophy, there are several different potential concerns to pursue. Here I will concentrate on two: the relevance objection and the no answer objection.

According to the relevance objection, while we are ignorant of certain types of fact, we are not ignorant of any type of fact that is *relevant* to consciousness. There are a number of different ways to develop this idea. Jerry Fodor, for example, in a review of Chomsky's *New Horizons,* argues that consciousness "seems to depend, not on the 'ultimate' nature of matter, but on its macrostructure" (Fodor, 2000, 3).[12] The point underlying this remark can, I believe, be stated as follows. Premise 1: all facts relevant to consciousness are macrostructural facts. Premise 2: there are no macrostructural facts of which we are ignorant. Conclusion: there are no facts relevant to consciousness of which we are ignorant.

This argument is certainly valid and we may agree with Premise 1. But someone who advances the ignorance hypothesis and the epistemic view will immediately deny Premise 2.

---

[11] For some discussion of these points, see Stoljar 2015 and Alter 2016, and the references therein.

[12] It is not clear that Fodor has consciousness specifically in mind here, but we may adapt what he says to this case.

After all, their whole point is that there are relevant facts of which we are ignorant. If so, this argument is no good against the epistemic view, since its second premise is no better than an assertion that the view is mistaken.[13]

Someone may try to motivate Premise 2 by pointing out (a) that there are *micro*structural facts of which we are ignorant—for example, facts associated with quantum mechanics—but (b) that this seems not to have an impact on macrostructural facts—for example, that quantum indeterminacy seems to be a feature of the micro-world rather than the macro-world. But even if these claims are true, they lend no plausibility to Premise 2, a point that comes out best if we consider the matter abstractly. Suppose we are ignorant of a particular type of fact, say A-facts. And suppose A-facts are irrelevant to a distinct type of fact, say B-facts. It scarcely follows that we are *not* ignorant of B-facts. In general, you don't get to know a type of fact simply by *not* knowing some quite distinct type of fact!

A different way to develop the relevance objection is suggested by David Lewis when he remarks, "the physical nature of ordinary matter under mild conditions is very well understood" (1999, 292). Lewis is here talking about the brain, and his point is that *its* physics is well understood, since it is an example of ordinary matter. This suggests the following version of the relevance objection. Premise 1: all facts relevant to consciousness are facts about the brain. Premise 2: there are no facts about the brain of which we are ignorant. Conclusion: there are no facts relevant to consciousness of which we are ignorant.

Once again the argument is valid, and we may agree with Premise 1. But why believe Premise 2? Of course *some* physical features of the brain are well understood—we know how its mass or shape, for example, is a function of the mass or shape of its parts. (It is these features of the brain that Lewis is talking about when he mentions "ordinary matter under mild conditions".) But other features of the brain—to take an obvious example, how exactly consciousness depends on it—are clearly not understood; indeed it is a scientific truism that this is so.[14]

---

[13] It is worth emphasis here that the ignorance hypothesis does not require that we are ignorant of facts concerning macrostructural *objects* or *first-order* macrostructural properties (i.e., properties of objects). It is enough for the thesis that we are ignorant of facts concerning *second*-order macrostructural properties, i.e. properties of first-order properties. For some discussion of this using a slightly different terminology, see Stoljar 2006, 72-74.

[14] One might object that the point in the text would be stronger if there were examples of a principled incompleteness in our theory of the brain or of the physical that's independent of consciousness. For a number of historical cases along these lines, see Stoljar 2006, 123-141.

One might try to bolster Lewis's version of the relevance objection by pointing out that we have a good understanding of what the individual parts or constituents of the brain are. At least since the work of Ramon y Cajal at the end of the 19th century, it has been known that the brain is made up of cells, which are themselves made up of molecules and so on and so forth (cf. Shepherd 1991). However, for the epistemic view the crucial issue is not what the individual parts of the brain are. It is rather what the properties of those parts are, and indeed what the properties of the properties of those parts are. In fact, it is precisely for this reason that Lewis makes his point by speaking of the 'physical nature' of the brain rather than its constituents.

A final way to develop the relevance objection appeals to the idea in philosophy of mind known as *functionalism*, which says, roughly, that mental states are identical to functional states.[15] The key thing about functional states, according to this line of thought, is that they make no (or at least very few) demands on the matter that makes them up; hence even if we are ignorant about underlying features of matter, we may know exactly what the functional states are. This suggests a third version of the relevance objection, as follows. Premise 1: all facts relevant to consciousness are functional facts, i.e., facts about functional states. Premise 2: there are no functional facts of which we are ignorant. Conclusion: there are no facts relevant to consciousness of which we are ignorant.

But the problem here is that the notion of a functional fact can be interpreted in many different ways. Formally speaking, functionalism is not so much a theory as a technique of definition; it says, for example, that a mental state M is the nth state in a sequence of states that satisfies the relevant description. Everything depends therefore on what the relevant description is. If there are no constraints on the description, Premise 1 of the above argument is certainly plausible, but Premise 2 is implausible, since even unknown and unknowable facts may be functional facts in this general sense. If there *are* constraints on the relevant description, however, we get the reverse problem. For example, in philosophy of mind, people who describe themselves as 'functionalists' typically assume that the descriptions at issue are limited to those concerning causal processes, which results in the idea that mental states are exhausted by their causal roles. This view renders Premise 1 plausible, but equally it renders Premise 2 *im*plausible, since it is implausible that all facts relevant to consciousness are functional facts in this restricted sense. Indeed, the idea that all facts relevant to

[15] For a very good presentation of functionalism, see Braddon Mitchell and Jackson 2007

consciousness are functional facts in this restricted sense is itself refuted by the conceivability argument. Either way, therefore, this version of the relevance objection is unpersuasive.

The No Answer Objection

For these reasons, the epistemic view is in a good position as regards the relevance objection. What then of the other objection I mentioned, that properly understood the view is no answer to the hard problem at all?

We may formulate this objection by looking back at the two questions I distinguished earlier. Question A was: what is the explanation of consciousness? Question B was: does consciousness have a reductive explanation? I suggested before that, while the epistemic view is focused on Question B, one may well insist that a genuine contribution to the hard problem should focus on Question A. And it is this observation that sets up the no answer objection, which we may state as follows. Premise 1: something is a genuine contribution to the hard problem of consciousness only if it answers Question A. Premise 2: the epistemic view does not answer Question A. Conclusion: the epistemic view is not a genuine contribution to the problem.

How to react? I think the weakness here comes to light when we focus on what exactly Question A is asking; that is, what question is at issue when people ask for 'the explanation of consciousness', and, connected to this, what an answer to that question might be. For while we sometimes take ourselves to know intuitively what an explanation of consciousness would be, this becomes much less clear when you focus on it.

One attractive way to approach this issue starts from the observation that consciousness is something that has both a history and a constitution.[16] As regards history, each and every conscious event stands at the end of a huge causal history leading all the way back to the beginning of the universe. As regards constitution, each and every conscious event stands at the apex of a huge constitutive hierarchy leading all the way down to the basic elements of the universe. One thing you could be asking for when you ask for 'the explanation of consciousness' is total information about both the history and constitution of consciousness. Hence an answer to Question A would be a proposition expressing this total information.

---

[16] This way of looking at things owes a lot to Lewis's discussion of causal explanation; see Lewis 1986. For further discussion see Stoljar forthcoming.

Now, if that is how Question A is to be understood, there is no doubt that the epistemic view does not answer it; that is, Premise 2 of the no answer objection is correct. Indeed, more than not answering it, the epistemic view entails that it cannot now be answered, since it entails that at least some of the relevant facts here are unknown to us.

By the same token, however, Premise 1 is on this interpretation very implausible, since it is very implausible that something is a genuine contribution to the hard problem only if it provides total information in this sense. For one thing, *no* contribution to the hard problem has *ever* provided total information; hence nothing *ever* said about the hard problem constitutes a contribution by that standard.

Moreover, many of the most interesting questions about consciousness can be pursued in the absence of total information. We have already noted that the hard problem is a cluster of related questions, and singled out Questions A and B as members of that cluster. Several further questions in the cluster are as follows:

Question C: What is it that makes a mental state conscious?
Question D: Are you always aware of your conscious states?
Question E: How do those states interact with other psychological features?
Question F: How do they evolve over time?
Question G: What is their epistemological and rational role?
Question H: In what ways, if any, are they valuable?
Question I: What neural and computational structures are associated with them?

We may profitably formulate and assess answers to these (and other) questions without having total information about the history and constitution of consciousness. Hence no answer to these questions would count as contributions to the hard problem either given the standard we have just set out.

If we interpret Question A as asking for total information about the constitution and history of consciousness, therefore, Premise 1 above should be rejected. Of course, one might say now that Question A should *not* be interpreted as asking for total information but rather for partial information, or better partial information of a contextually relevant sort. One might well say this, and then Premise 1 is plausible. But, unfortunately for the proponent of the no answer objection, now Premise 2 is *im*plausible! For the epistemic view *does* provide partial information of this sort. If they are sound, the famous arguments in philosophy of mind about zombies, Mary, inverted spectra and the rest will tell us something extremely

dramatic, namely, that consciousness marks a major constitutional and historical break in the system of nature. The suggestion of the epistemic view is that we can reject those arguments so long as we agree that we are ignorant of certain features of the situation—something that is highly plausible in any case.

Conclusion

My main aim has been to set out the elements of the epistemic approach to the problem of consciousness, and to respond to two objections. Obviously there are further objections to consider; I will not take them up here. Let me instead end by asking why the epistemic approach is not more common than it is.

In saying this, I am not denying the view has had defenders; on the contrary, we noted a number of extremely distinguished ones at the outset. What I mean rather is that the bulk of contemporary philosophers and scientists of mind these days give it short shrift, or at least so it seems to me. Why so?

Partly the answer is that people are not clear enough on what the view is, on what questions it is trying to answer and, more important, what questions it is not trying to answer. For example, one common mistake is that the proponent of the epistemic view is no better than Bob who we met before. But I think there are also deeper issues at work here, issues having to do with some of the dominant ways we have had in the last hundred years of framing and pursuing philosophical problems.

It is not too much of an exaggeration to say that philosophy of mind, and philosophy in general, in the last hundred years has had two main phases: a positivist phase and a post-positivist phase. The positivist phase is typified by Carnap and has been recently revived and reimagined in interesting ways by David Chalmers (see Carnap 1967, Chalmers 2012). The main idea of this approach is to assume, first, that there is a base language whose constituent expressions we understand and, second, that every truth about the world bears the right kind of relation to a truth that can be formulated in this base language—base truths, as we may call them for short. Obviously, views of this kind may differ on what exactly the base language is, and what relation must obtain between all truths and base truths.

The post-positivist phase is typified by Quine and is widely represented in contemporary philosophy (see, e.g., Quine 1953, 1960). The main idea of this approach is to assume, first, that we have a set of base facts given to us by contemporary science and, second, that something is a fact only if it bears the right sort of relation to one of these base

facts. Obviously again, views of this kind may differ on what exactly the base facts are, and what relation other facts must bear to these base facts.

On the surface there is a lot that differentiates the Carnapian and the Quinean. The Carnapian view lends itself to the formal mode, the Quinean view to the material mode. Quineans tend to interpret themselves as privileging science over philosophy and criticize Carnapians for doing the opposite. Carnapians tend to be more sympathetic to conceptual analysis and the a priori than Quineans.

But these differences should not blind us to the fact that there is also something deeply similar here. Both views tend to assume that philosophical problems take the form of explaining various problematic sentences or facts by relating them in acceptable ways to sentences or facts of an approved sort. This sets up a very strong expectation about what a contribution to the hard problem of consciousness will look like: in the Carnapian case, it will be an account of how sentences about consciousness relate to sentences in the base language which we understand; in the Quinean case, it will be an account of how facts about consciousness relate to facts presented to us in contemporary science which again we understand.

I think there is no doubt that the epistemic view of the hard problem violates expectations of this sort. It does not 'explain' consciousness in this sense. However, this is more a problem for these expectations than for the epistemic view. Certainly we would like an explanation of consciousness in the sense mentioned in the previous section: the provision of contextually relevant information about the history and constitution of consciousness. But nothing here requires that the information in question be provided in ways recommended either by Carnap or Quine. A more plausible view is that this information will be provided over the long term by a combined effort of various different disciplines and enterprises. Philosophy of mind is one of these disciplines, and an important one, but not the only one.

It is at points like these that we arrive at one of the most interesting aspects of the epistemic approach to the hard problem of consciousness. Not only is the view historically important, and extremely plausible from a scientific and philosophical point of view. It also represents a way of doing philosophy that is quite different from the ways we have settled into in the last hundred years.

References

Alter, T. 2015 'The Structure and Dynamics Argument Against Materialism', *Nous* (2015) 1–22

Alter, T and Howell, R (eds.) 2012 *Consciousness and the Mind-Body Problem: A Reader* (Oxford: Oxford University Press)

Arnauld, A. 1641 'Fourth Set of Objections to Descartes' *Meditations*' in J.Cottingham et. al. (eds.) 1985b *The Philosophical Writings of Rene Descartes*, Volume II Cambridge: Cambridge University Press

Braddon Mitchell, D and Jackson, F 2007 *Philosophy of Mind and Cognition*, 2nd Ed, Wiley-Blackwell

Du Bois-Reymond, Emil. 1886 *Reden von Emil Du-Bois-Reymond* [Addresses of Emil du Bois-Reymond], (Leipzig: Erste Folge, Verlag von Veit & Comp)

Du Bois Reymond, E. 1872 'The Nature of Scientific Knowledge'. In Du Bois-Reymond 1886

Carnap, R 1967 *The Logical Structure of the World,* University of California Press (Original publication in German: 1928)

Chalmers, D 2012 *Constructing the World* Oxford: Oxford University Press

Chalmers, D. and Jackson, F. 2001 'Conceptual Analysis and Reductive Explanation' *Philosophical Review* 110 (3): 315-61

Chomsky, N. 1975 *Reflections on Language* London: Pantheon Books

Chomsky, N. 1986 *Knowledge of Language: Its Nature, Origin and Use* New York: Praeger Special Studies

Chomsky, N. 2000 *New Horizons in the Study of Mind and Language* Cambridge: Cambridge University Press

Chomsky, N. 2009 'The Mysteries of Nature: How Deeply Hidden?' *Journal of Philosophy*, 106 (4): 167-200

Demircioglu, E. 2016 'Human Cognitive Closure and Mysterianism: Reply to Kriegel' *Acta Analytica*, 1–8

Descartes, R. 1985b *Meditations on First Philosophy* in Cottingham, J. et. al. (eds.) *The Philosophical Writings of Rene Descartes* Volume II Cambridge: Cambridge University Press Original date: 1641

Flanagan, O. 1992 *Consciousness Reconsidered*, Cambridge, MA: MIT Press.

Fodor, J. 2000 Review of Chomsky's *New Horizons in the Study of Language and Mind*, Times Literary Supplement

Jackson, F. 1982 'Epiphenomenal Qualia', *Philosophical Quarterly,* 32, 127-36 Reprinted In Ludlow et al 2004; references to the reprinted version

Kriegel, U. 2003 The new mysterianism and the thesis of cognitive closure. *Acta Analytica*, 18(30/31), 177–191.

Lewis, D. 1986 'Causal Explanation' in Lewis, D 1986, *Philosophical Papers,Vol. II* New York and Oxford: Oxford University Press, 241-240

Lewis, D. 1994 'Reduction of Mind' in Guttenplan, S (ed.) 1994) *A Companion to Philosophy of Mind* London: Blackwell, reprinted in his *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press, 1999), 291-324 All references are to the reprinted version.

Locke, J. 1975 *An Essay Concerning Human Understanding* ed. P.H.Nidditch Oxford: Oxford University Press

McGinn, C. 1989. Can We Solve the Mind-Body Problem? *Mind*, 98: 349–366

McGinn, C. 1991. *The Problem of Consciousness* (Cambridge: Blackwell)

McGinn, C. 2004. *Consciousness and Its Objects* (Oxford: Clarendon Press)

Nagel, T. 1974 What Is It Like to Be a Bat? *Philosophical Review*, 83: 435–450

Pinker, S 2007 The Brain: The Mystery of Consciousness Time, available at: http://content.time.com/time/magazine/article/0,9171,1580394-6,00.html

Priestley, J 1977 'Of the Properties of Matter', in Passmore, J (ed.) 1965 *Priestley's Writings on Philosophy, Science and Politics* New York: Collier Books, pp.103-7

Russell, B. 1927 *The Analysis of Matter* (London: Kegan Paul)

Shepherd, G. M. 1991 *Foundations of the neuron doctrine* Oxford University Press

Stoljar, D, 2005 Review of McGinn's *Consciousness and Its Objects, Notre Dame Philosophical Review*s, 2005

Stoljar, D, 2006 *Ignorance and Imagination* New York: Oxford University Press

Stoljar, D. 2010 *Physicalism* London and New York: Routledge

Stoljar, D. 2015 Russellian Monism or Nagelian Monism? In T. Alter and Y. Nagasawa (eds.) *Consciousness in the Physical World: Perspectives on Russellian Monism.* New York: Oxford University Press: 324-345.

Stoljar, D forthcoming *Philosophical Progress* Oxford: Oxford University Press

Wittgenstein, L. 1960 *The Blue and Brown Books* Oxford: Basil Blackwell