

# Consciousness, Materialism, and Ideology

William S. Robinson<sup>1</sup>

I

I am honored to have the pleasure of presenting a few ideas to the Advanced Reasoning Forum for the first time. In view of the name of this group, I believe it will not be out of order for me to begin by discussing a piece of advanced reasoning—that is to say, a piece of reasoning that was on the forefront of thought when it was put forward about 370 years ago. I should warn you that it is somewhat controversial to say that the argument I am about to discuss was offered by Descartes, for it is not perfectly transparent that he argued in precisely the way I am going to represent him as having argued. Nonetheless, the argument I will ask you to consider is fashioned from clearly Cartesian premises that occur very close together in the *Meditations*, and reflection on it will take us directly to the heart of the debate between dualism and materialism.

Let us briefly remind ourselves that by the middle of the second of his *Meditations*, Descartes has convinced himself of two things. One is that he exists; this proposition is safe even if there is a supremely intelligent and powerful Evil Demon who is determined to mislead poor Descartes. The other is that, with respect to the whole of the world of bodies, he is in danger of deception. The Evil Demon might present Descartes with the appearances of bodily things, even if there is nowhere anything that corresponds to those appearances, in the way that Descartes has been accustomed to believe that there are corresponding bodily things.

Having convinced himself of this much, Descartes is in a position to argue as follows.

- (D1) I am certain that I exist.
- (D2) I am not certain of any bodily thing that it exists.
- (D3) I cannot be both certain and uncertain of the (existence of the) same thing at the same time.

*Therefore,*

- (D4) I am not (identical with) any bodily thing.

In view of the fact that the word “certain” occurs in all the premises of this argument, it is somewhat natural to expect to find it in the conclusion. This naturalness may lead you to want to reformulate the conclusion of the argument, for example, to “I am not certain that I am identical with a bodily thing.” I have no objection to anyone’s considering the merits of such revised arguments, but I regard them as less interesting than the one stated. In any case, even if you think

<sup>1</sup> The ideas below were originally discussed in a seminar that I presented to the Advanced Reasoning Forum in Humor, Romania, in July, 2000. For the sake of readability, I have written the following as a lecture, but many of the points made were developed cooperatively in the seminar.

some revised argument would be a *better* argument, there can be no objection to considering the merits or demerits of the argument that I have actually stated; and it is this task that I ask you to undertake with me. It will, perhaps, increase the interest of this task if I remind you that Descartes drew the strong conclusion—that he is distinct from any bodily thing—and not merely that he wasn't certain he was a bodily thing.<sup>2</sup>

The third premise of the foregoing argument (let us call it “the *D*-argument”) is perhaps the least transparent sentence in the argument. Let us therefore consider a second version of essentially the same argument, which will make the force of (D3) a little clearer. This “*D'*-argument” is stated in terms of a somewhat peculiar property, which I will abbreviate by “*C*”. *C* is the property that a thing, *x*, has, just in case *x* is known (with certainty) by me to exist.

- (D'1) I have property *C*.
- (D'2) No bodily thing has property *C*.
- (D'3) For all *x* and *y*,  $x = y$  if and only if  
for all properties *P*, *Px* if and only if *Py*.  
(Leibniz's Law for identity.)

*Therefore,*

- (D'4) I am not identical with any bodily thing.

Well, now, what shall we say? Are we in possession of a good argument for the conclusion that I am not identical with any bodily thing? Few philosophers indeed would be willing to say that either the *D*- or the *D'*- argument gives us a good argument for its dualistic conclusion. The reason is the standard one: there are plainly invalid arguments that have the same form as the foregoing arguments. We may, for example, imagine a detective who is investigating a theft and who is interrogating a suspect, John Jones. Let us imagine this detective to reason as follows.

- (1) John Jones has the property of being certainly known by me to be sitting before me.
- (2) The thief (in the case under investigation) does not have the property of being certainly known by me to be sitting before me.
- (3) Leibniz's Law for identity is true.

*Therefore,*

- (4) John Jones is not (identical with) the thief.

Here, surely, is a candidate for quick demotion.

Someone might object that this argument concerns only accidental properties, and that Descartes believed he was considering an essential (necessary) property of himself. But necessity makes no difference, as we can see from the following example of a beginner at geometry.

<sup>2</sup> Strictly, Descartes did not draw his conclusion until he thought he was in a position to know that the goodness of God certified the correctness of his reasoning powers. Nonetheless, once he believed he had such assurance, it was the foregoing argument that led him to the claim of distinctness of himself (or, thinking substance) from anything bodily (extended substance).

- (1) I am certain that I am thinking of a right angled triangle.
- (2) I am not certain that I am thinking of a figure that has the Pythagorean property.<sup>3</sup>
- (3) I cannot be certain and uncertain of the same thing at the same time.

Therefore,

- (4) The figure of which I am thinking does not have the Pythagorean property.<sup>4</sup>

A little reflection will lead us to a generalization of the difficulty in these arguments. There is nothing very special about "certain". We will get parallel invalidities with "know" and "believe", and even with "hope for" and "desire". The presence of a psychological verb (and certain other contexts) defeats the D- and D'-arguments: We cannot apply Leibniz's Law (without special provisos) in contexts where psychological verbs are present.<sup>5</sup>

## II

It has taken us a few minutes to arrive at the generalized reason for the failure of the D- and D'-arguments, but it was clear almost immediately that there was something wrong with them.<sup>6</sup> The ease of seeing that there is something amiss

---

<sup>3</sup> The Pythagorean property is the property of having the square of one side equal to the sum of the squares of the remaining sides.

<sup>4</sup> In D' form:

- (1) A right angled triangle has the property of being certainly known by me to be that of which I am thinking.
- (2) No figure that has the Pythagorean property has the property of being certainly known by me to be that of which I am thinking.
- (3) Leibniz's Law for identity is true.

Therefore,

- (4) The right angled triangle of which I am thinking does not have the Pythagorean property.

<sup>5</sup> Chisholm in *Perceiving: A Philosophical Study*, tried to use (apparent) failure of Leibniz's Law as a mark of the mental. However, the problem also arises in contexts containing modality or quotation. Following Quine, *Word and Object*, we may note that nine is necessarily identical with nine, but nine is not necessarily identical with the number of the Sun's planets; we cannot validly conclude that nine is not the number of the Sun's planets. We may further observe that "eerie" has five letters. "The sixth word of the fourth sentence of this note" does not have five letters. We cannot validly conclude that the sixth word of the fourth sentence of this note is not "eerie".—The remainder of this lecture, however, will be concerned only with cases involving psychological verbs.

What do I mean by "without special provisos"? It is open to us to insist as, e.g., Quine does in *Word and Object*, that we preserve Leibniz's Law by adopting referential readings. On such a reading, our detective is simply not entitled to premise (2); i.e., regarding the thief, it may well be true that the detective certainly knows that he is sitting before him. (He just doesn't recognize this truth when the person is described as "the thief" rather than as "John Jones".) But such readings often do not come naturally; hence, it is a "special proviso" that we insist on requiring them.

<sup>6</sup> Several members of the seminar were ready to offer plainly invalid parallels of their own before I had finished stating the D'-version.

ought to make us a little suspicious. Indeed, one plausible reaction would be that Descartes cannot really be supposed to have advanced an argument that can so easily be seen to be invalid. I am inclined to agree, but the argument seems to me so close to what Descartes says that I do not think we can reject it outright. Instead, I suspect that Descartes was tacitly supposing something that made the D-argument look better to him. I am thus led to look for a premise that we might add to the D-argument (or to some slightly revised version of it) that will significantly improve the argument. Such an additional premise should, of course, be one that Descartes himself could reasonably be supposed to have accepted, even if he never made it explicit.

Now there is an idea that comes up from time to time in contemporary discussions of phenomenal consciousness that will provide us with a candidate for Descartes' missing premise. I would not say that this is a point that is often emphasized, but most philosophers will at least recognize it as something that Descartes' contemporary descendants are likely to hold. (These descendants are qualia realists, also known as "qualia freaks".) Not to tantalize you, the idea can be baldly put in this way: There is nothing to qualia (or to phenomena, experiences, raw feels, sensory consciousness, pains, afterimages, etc.) other than the properties that constitute them as the qualia that they are. There is no inner or hidden nature to them; their presented surface is their nature, and their entire nature is to be that presented surface.

Let us approach this key idea a little less baldly but perhaps with more understanding. Consider a plain fact, a fact that is the root of much philosophy, a fact with which Descartes was intimately (some might say excessively) concerned, namely: Things can appear other than they are, and can even appear when they aren't there at all. What is the logic behind this familiar fact? First, there is a contrast between a *thing* and *the way it appears*. ("Appears" may seem a little formal; that is because it is most often replaced by one of its species, "looks", "sounds", "tastes", etc.) An ordinary thing, e.g., a table or a tree, is, in general, a complex object. Even when it appears the way it is, it is understood that a thing can have aspects that do not appear: insides, backsides, microstructures, and relations to other things. And, of course, for any number of reasons, a thing may *not* really be as it appears to be. Now, in this statement, the way things appear (or, the appearance of a thing) has a certain role, which is very limited. The way a thing appears is just the way that thing presents itself; an appearance has no other job than to be the way a thing appears. Serving that function, so to speak, exhausts its role.

Now, nonassignment of a further role need not imply inability to have a further role. But let us see what it would be like to embrace "thick" appearances, that is, to regard appearances as possessors of aspects other than that of merely being the way a thing appears. The moment we enrich our conception of appearances in this way, we will find it necessary to say that not all of their aspects appear (to their possessors). And the moment we say this, we will be committed to introducing a new level of appearances: we will have to have our

old appearances of things, and then, in addition, the way those old appearances happen to be appearing to their possessors. We will be so committed, because we will have to have some way of enforcing a distinction between the aspects of (our old) appearances that are the way they are appearing, and the aspects of (our old) appearances that are not the way they are appearing.

I hope your suspicions are thoroughly aroused by this talk of appearances of appearances, or the way appearances appear. For one thing, the grammar of things and their appearances, or things and the ways they may appear, is very familiar; but talk of appearances appearing, or the way appearances appear is not. Behind this grammatical observation is a more serious problem. If the new appearances are *appearances*, that is, appearances in the same sense as the old appearances, there would seem to be no reason why they too may not be “thick”, i.e., no reason why they may not have aspects that are not apparent; and then we will need the way appearances of appearances appear, and so on. While such a stack of appearances is not a *vicious* regress, the lack of a nonarbitrary place to stop makes it a theoretical monstrosity. The natural response of any theorist will be to cut the regress off at the earliest possible step. Since we cannot get along with no appearances at all, the best we can do is to refuse to allow appearances of appearances. In that case, we must regard appearances as *not* being possessors of aspects that do not appear. We must say, instead, that there is nothing to an appearance other than its surface; it just *is* the aspect that some *thing* appears *as*. Appearances don’t appear, they *are*; and *their* being is just some *thing’s* appearing.

Although I know of no text in which Descartes says anything like the foregoing, the problem of allowing appearances to be “thick”, that is, to have aspects that do not appear, is so obvious that it is plausible to imagine that Descartes thought of appearances as “thin”, that is, as being exhausted by the property that they present something (some thing) as having. “Thin” appearances also seem to be the outcome of Descartes’ method of doubt. One can doubt that reality is appearing to one as it is, but it is much harder to doubt that the appearances are what they are. Alleged nonapparent aspects of appearances, however, would have to go on the side of realities “behind” the appearances, and would thus fall within the realm of the dubitable. If, as Descartes seems to do, one confines appearances to what is indubitable, then appearances must not have “hidden” aspects, i.e., they must be “thin” in the sense we have been using this term.<sup>7</sup>

If this point is found agreeable, the next question will be how the thinness of appearances might be of any help with the D-argument we reviewed earlier. Here

<sup>7</sup> Descartes says in *Meditation II*, Haldane and Ross (1931), translators:

Finally, I am the same who feels, that is to say, who perceives certain things, as by the organs of sense, since in truth I see light, I hear noise, I feel heat. But it will be said that these phenomena are false and that I am dreaming. Let it be so; still it is at least quite certain that it seems to me that I see light, that I hear noise and that I feel heat. That cannot be false; properly speaking it is what is in me called feeling; and used in this precise sense that is no other thing than thinking.

again, I shall offer a rough and very bald formulation. After that will come some refinement, and then, at the last, an evaluation of the newly-formulated argument.

The rough formulation is this. Let us suppose that Descartes is thinking that what he is certain of is the existence of *appearances*. Now, *appearances* are thin: They don't have aspects that don't appear. So, if there is some property that is not appearing to Descartes, when he does have an appearance, that property is definitely not a property of the appearance. And if there is some property, *P*, that is never a property *in* any appearance, then that property is definitely not a property *of* any appearance, and appearances can be taken to be distinct from anything that has *P*.

Before I set this argument out a little more formally, let us think for a moment about our incompetent detective's reasoning about the thief. The way this incompetence works is this. The thief is a thing; it is "thick". It is the kind of thing that can have properties that are "hidden", i.e., not apparent. Another terminology in which this kind of point is sometimes made derives from Frege. One may say that a thief may be presented under one mode of presentation (e.g., as the person sitting across the table from me) without being presented under another mode of presentation (e.g., as taker of jewels).<sup>8</sup> Our detective's mistake may be described as an invalid inference from "is not presented under mode of presentation *M*" to "is not *M*". But, just as appearances do not plausibly appear, so modes of presentation do not plausibly have modes of presentation. Thus, if Descartes were thinking of his argumentation as applying only to modes of presentation, then the detective counterexample would violate a restriction on the universe of discourse. It would thus be irrelevant to Descartes' argument, and his reasoning might not be incorrect after all. The same point can be made about our geometrical example. Abstract though they may be, geometrical figures (and other mathematical entities) can be presented in more than one way. They are not mere modes of presentation; so, if Descartes is restricting his argument to modes of presentation, the geometrical example is also irrelevant to his case.

Let us now see if we can incorporate these suggestive remarks into a formal statement of a neo-Cartesian argument.

- (ND1) I am certain I am having an appearance that is constituted by the property *F*.
- (ND2) I am not certain that anything has the property *G*. In particular, *G* is not a property that is constitutive of any appearance I (now) have.
- (ND3) Appearances have no properties other than those that constitute them.  
*Therefore,*
- (ND4) The appearance I am having does not have the property *G*.

Several comments are in order. "Constituted by" is a new term, and so needs some explication. Let us start with the basic forms of appearance

<sup>8</sup> Likewise, a figure may be presented under the mode of presentation "right angled triangle" without being presented under the mode of presentation "bearer of the Pythagorean property". A number may be presented as odd without being presented as prime, or as the eighth term of the Fibonacci series, etc.

statements: “Object *O* appears to subject *S* as *F*” and “It appears to *S* that there is an *F* object.” In these forms, “*F*” tells us what kind of appearance *S* is having. To say that an appearance is constituted by a certain property is just to identify that property as the one that substitutes for “*F*” in these forms.

On this understanding, the ND-argument has a strong claim to be regarded as valid. This claim should be greeted with initial suspicion, because the psychological predicate “certain” occurs in the first two premises. I have included this term so as to make the connection to Descartes’ text visible, but now we can note that its contribution to the ND-argument is quite different from the contribution that “certain” makes to the D- and D’-arguments. In (ND1), all that “certain” does is to emphasize the epistemic status of an assertion that does the real work in the argument, namely, “I am having an appearance that is constituted by the property *F*.” Likewise, the argumentative force of (ND2) lies in the claim that I am not having an appearance that is constituted by the property *G*. To emphasize the point being made here, let us reformulate the argument a little.

(RND1) I am certain I am having an appearance that is constituted by the property *F*.

(RND2) I am not certain that I am having an appearance that is constituted by the property *G*.

So,

(RND3) I have an appearance that is constituted by the property *F*.

(RND4) I do not have an appearance that is constituted by the property *G*.

(RND5) Appearances have no properties other than those that constitute them.

Therefore,

(RND6) The appearance I am having does not have property *G*.

Whatever force this argument has remains (or, perhaps, increases) if we lop off the first two premises. These premises somewhat redundantly say that I am warranted in asserting (RND3) and (RND4).

(RND6) is considerably weaker than Descartes’ intended conclusion.

However, suppose there is a property *G* such that

(RND7) I can never have an appearance that is constituted by the property *G*.

Then we may derive the stronger conclusion:

(RND8) No appearance has the property *G*.

Descartes’ candidate for a property *G* that would make (RND7) true is *extension*. There are problems with this candidate. On the one hand, visual experience, tactile experience, and bodily sensations all have a “spread out” character to them. Sounds also have a locational aspect. If “extension” is understood in such a way as to conflict with these points, then its use simply falsifies the phenomenological facts.<sup>9</sup> And on the other hand, “extension”, as

<sup>9</sup> This point is further explained and supported in my book in progress, *Understanding Phenomenal Consciousness*; and in Chapter 6 of M. Velmans’ *Understanding Consciousness*.

Descartes understood it, does not include mass. It is therefore inadequate to capture what contemporaries would understand to be the *physical* world.

The failure of Descartes' candidate for a property that could serve as *G* in (RND7) does not, however, show that there is no good candidate, and in fact there is a better one. None of our appearances appear to us as neurophysiological events (e.g., ratios of neural firing rates, or rises of activations in particular brain parts). Thus, if the RND-argument is accepted, and neural event properties are substituted for "*G*", we will have an argument that our appearances are not neural events.<sup>10</sup>

This result is a strong neo-Cartesian result, but still not as strong as Descartes' own which, let us recall, was that I am not any bodily thing. I hope no one will be disappointed to find out that I am not going to defend Descartes' putting his conclusion in this way. Despite Descartes' "modernism", that is, his break with Scholasticism, he was so habituated to thinking in terms of substances and accidents that he was unable to avoid an unsupported substantialist outlook. However, disavowing Descartes' substantialism should not be regarded as a reason for rejecting the view that the RND-argument is genuinely neo-Cartesian. There are many passages in which Descartes limits what should be thought to "pertain" to "himself", by excluding everything of which he can doubt, that is, everything that is not the way he is thinking of something or the way something seems to be appearing to him. To put the matter in an intentionally paradoxical way, Descartes argues repeatedly that

I am nothing but the ways I think and the ways it seems things appear to me.

In contrast, the upshot of accepting the RND-argument is only that appearances are not identical with neural events, or with any physical event of type *G* such that nothing appears as *G*.

Substantialism about selves is not the only kind of ontological worry one might have about the RND-argument. "Appearance(s)" functions as a noun in this argument or, as one may also say, the argument treats appearances as *things* (albeit in a very inclusive sense of "thing"). But perhaps appearances are better thought of as *events*; and perhaps it is not innocent to apply a logic of things to events.<sup>11</sup> So, it may be that a better understanding of event-logic would bring to the surface some now-hidden reason for doubting the validity of the RND-argument.

I have no decisive reply to this objection, but I am not very worried by it. While I am no specialist in the logic of adverbs, I am under the impression that there is no generally accepted account of how to treat adverbial inferences. This fact does not make me doubt that "Juney barks loudly, therefore, Juney barks" is valid. I suspect that there are countless similar arguments that seem valid, and

<sup>10</sup> There is nothing in this conclusion that conflicts with the view that neural events *cause* appearances, and this is, indeed, the view of qualia realists.

<sup>11</sup> This objection reflects my understanding of a point that was raised in the Humor seminar by Richard Epstein.



that will continue to be regarded as valid whenever a satisfactory account of inferences involving adverbial constructions is developed. The RND-argument is fairly simple, and the key concept, appearance, is well understood. Thus, I suspect that the apparent validity of the RND-argument will survive its recasting into a properly articulated event logic, whenever such a logic is developed.

### III

I have tried to explain the RND-argument in such a way that it will recommend itself as an important argument for dualism (or, to use more contemporary language, anti-physicalism) with respect to phenomenal qualities (appearances, qualia, sensory experience, phenomenal consciousness, etc.). But I know that contemporary materialists will not be moved by it. One might suppose that the reason is that they will doubt the new premise,

(RND5) Appearances have no properties other than those that constitute them.

But doubting this premise would not serve the purpose of materialists. Their view is not merely that all *things* are physical, that is, that all things have at least one physical property, whatever other properties they may have. Their view is that all *properties* are physical properties. What is necessary in order to maintain this view is to hold that the RND-argument is *invalid*, and that it is so because

(M) The property *F* and the property *G* can be identical properties, even if *F* is constitutive of an appearance and *G* is not constitutive of that appearance.

If (M) is true, then the premises of the RND-argument can be true while its conclusion is false.

In this lecture, I cannot survey all of the recent literature produced by materialists and their opponents. The most I can do is to illustrate some of the points I will make by showing how they occur in the work of one materialist author, and I will do this shortly. But the conclusion I have drawn from an extensive involvement with this issue is that, after many layers of surface issues have been resolved or dissolved, the irreducible and irremovable core of the contemporary materialist/anti-materialist debate is one's acceptance or rejection of (M). My further remarks in this lecture will thus concern the prospects for establishing or refuting this key proposition.

I will begin by noting that (M) is certainly not obvious on its face, and seems to cry out for an explanation. "How," we may wonder, "could it happen that *F* is the way something appears, and *G* is not the way anything appears, and yet *F* and *G* are the very same property?" The matter becomes even more mysterious when we consider what materialists routinely offer as the substituends for "*F*" and "*G*". There are a number of proposals here, and we cannot go into the reasons for their differences. But a typical idea is that *F* may be a color—e.g., red may be a way something is appearing to us—and that very property may be the same property as a neural property, e.g., a set of ratios of activation potential rates in a collection of neurons.

It would seem that if such an identity is possible, we ought to understand how it is possible. But here we come upon one way of expressing the problem summarized in J. Levine's famous phrase, "the explanatory gap".<sup>12</sup> Namely, we have no explanation of how such property identities are possible, and we have no idea of what sort of research program might lead to such an explanation.

It is true that we have models for establishing property identities in some kinds of cases. Thus, for example, it is a result of chemistry that the property of being water is identical with the property of being H<sub>2</sub>O. Perhaps not everyone will agree with this result, on the ground that ice is H<sub>2</sub>O, but not water, or on the ground that a molecule of H<sub>2</sub>O isolated in interstellar space is not water. I think we should opt for theoretical simplicity and accept what the chemists say in this case, but we need not insist on this resolution. For, suppose we say that the property of being water is identical with the property of being a sizeable collection of H<sub>2</sub>O molecules within a certain temperature range. Then we will still be recognizing a case of property identity that is not immediately obvious, i.e., one that had to be discovered by scientific methods; and this case can serve as an illustration of how we can sometimes establish property identities.

Cases of property identity of this kind work by construction. That is, chemistry shows us how we can construct a body of water, namely, by assembling molecules of a certain kind. If anyone doubts this, there are many further explanations available to back it up. For example, water is transparent. It is so, because it does not absorb light in the visible frequency range; and this is true, in turn, because the resonant frequency of H<sub>2</sub>O molecules lies outside the visible frequency range. Again, water dissolves many substances, and does not dissolve others; these facts can be accounted for by the ways in which H<sub>2</sub>O molecules distribute themselves as a result of the forces exerted on them by molecules of other substances.

Now, the problem of the explanatory gap is that no explanations of this kind are available for understanding how, e.g., *red* could be "constructed" from neural events (ratios of neural event rates, etc.). Association, correlation, and causation present no problem. You can prick yourself with a pin, which is plainly a physical event, and produce a pain every time. Medical science shows us that pinpricks cause brain events, and that if we use anesthetics to prevent the usual brain events, we forestall the pains. So the *correlation* of pains and brain events is not in question and, since the pinpricks come first, the causation of pains by physical events seems questionable only on grounds that are common to causal cases in general. But neither correlation nor causation is identity, and the firmness of the correlational and causative claims contributes not one iota to *explanation* of how the property of being painful could be the *same property* as a property of a set of neural events.<sup>13</sup>

<sup>12</sup> "Materialism and qualia: the explanatory gap."

<sup>13</sup> Some philosophers, e.g., W. G. Lycan, *Consciousness*, will insist that the point be put as a claim of identity of having a pain with having a brain event of a certain kind. I will not go into the niceties of proper reformulation here. When these are carried out, however, the explanatory gap remains. See my book in progress, *Understanding Phenomenal Consciousness* for details.

The upshot of these remarks on explanation is that those who assert (M) are asserting a substantive claim that is nonobvious and requires explanation, but for which they have no explanation, nor even a plausible program for discovering the required explanation.

Now, it might be thought that those who deny (M) can turn these ruminations into a disproof of (M). Let us try to imagine how such an Anti-Materialist argument might go. I will put the idea in a general form, but for concreteness we may imagine that *F* is some color property, while *G* is some property of a set of neural events.

(AM1) *F* has the (second level) property of constituting an appearance.

(AM2) *G* does not have the property of constituting an appearance (for any appearance whatsoever).

(AM3) Two properties are identical if and only if they have all their second level properties in common.

(Leibniz's Law for property identity.)

Therefore,

(AM4) *F* is not identical with *G*.

Now, if our discussion in the first part of this lecture has been successful, you will be immediately suspicious of this argument. In fact, you will see just what a proponent of (M) will be likely to say. Namely, "constituting an appearance" is a psychological context, and we know that we are not entitled to make use of Leibniz's Law in psychological contexts. The AM-argument is thus invalid and is powerless to overthrow (M).

Actually, this formulation somewhat overstates what the materialist is entitled to. The contexts we can confidently affirm to vitiate applications of Leibniz's Law are contexts expressing propositional attitudes (e.g., beliefs or desires). What we are sure of is that we can believe or know about a thing under one mode of presentation, without knowing or believing anything about it under a different mode of presentation; and we can desire something under one mode of presentation without desiring it under another mode of presentation. But constituting an appearance is not constituting a propositional attitude. So, it *may be* that the context in the AM-argument is *not* one that vitiates applications of Leibniz's Law. Certainly, qualia realists will think of appearances as realities that do not count as propositional attitudes and that do not present the same kind of difficulty as the arguments that involve certainty, knowledge, belief or desire.

But materialists will not be contradicting themselves or begging the question if they say that the context in the AM-argument is one that *may* vitiate the application of Leibniz's Law. They cannot explain *how* it could do so, but they know that *some* contexts vitiate Leibniz's Law, and that some of these are psychological. Thus, they are entitled to say that the AM-argument *may use* a context that vitiates applications of Leibniz's Law, and that if it does, it is not a refutation of the key proposition, (M).

Putting together the points of this section leads to the following conclusion.

Materialism has not been refuted; we have found no way in which (M) can be *disproved*. But the way in which it can be rescued from refutation *requires* that materialism be *empty*. Materialists cannot explain why the AM-argument doesn't work, they can only make the negative point that a certain possibility has not been ruled out. They do not actually have an argument to show that the possibility (that AM is invalid) is a genuine possibility; they have only the right to say that the genuineness of this possibility has not been disproved. They have no explanation as to *how* (M) could be true (this lacuna is the explanatory gap); the only comfort they may find is the fact that (M) fails to be refuted.

The result that a defensible materialism is empty in the sense just explained is the most important conclusion of this lecture. Rather than just leave it to your consideration, however, I want to make it a little more memorable by giving a concrete illustration of some of the points that have led to it. Then, I want to use our conclusion and our illustration to raise a general, and somewhat disturbing, question about philosophical reasoning.

#### IV

Just so you won't think I have merely invented the dialectic of the previous sections, let us quote a little from a well known contemporary materialist, Patricia Churchland.<sup>14</sup> I will be considering her response to a view that I hope you will remember from our foregoing discussion: the view, namely, that "in the case of consciousness, the appearance is the reality. . . . Feeling the pain is all the reality there is to pain."<sup>15</sup> Her response to this view is:

What is troublesome is the idea that all the reality there is to a sensation is available through sheerly having it. How could you possibly know that? I suggest, instead, a rather simple alternative: A sensation of pain is real, but not *everything* about the nature of pain is revealed in introspection – its neural substrate, for example, is not so revealed.<sup>16</sup>

This passage introduces the idea that there is something about the "nature of pain"—not merely what is correlated with or causative of it—that is not revealed in introspection (or, in my terminology, is not constitutive of a kind of appearance). It is natural to ask what this hidden aspect of the nature of pain is. "Neural substrate" gives some hint of an answer, but in the following paragraph Churchland makes explicit what she has in mind.

Commonly, science discovers ways of taking us beyond manifest or superficial properties of some phenomenon. Light is refracted in water – that is observable. Light turns out to be electromagnetic radiation, a property not straightforwardly observable.<sup>17</sup>

The property of being electromagnetic radiation and the property of being refracted in water are here treated as distinct properties. But light is held to be

<sup>14</sup> P. S. Churchland, "Brainshy: nonneural theories of conscious experience".

<sup>15</sup> *Ibid.*, p. 117.

<sup>16</sup> *Ibid.*, p. 117.

<sup>17</sup> *Ibid.*, pp. 117–118.

electromagnetic radiation, in which case the property of being light and the property of being electromagnetic radiation are the same property. So, Churchland's model really is one of discovering identities. The same point holds of the other reductions she mentions as examples, namely, temperature to mean kinetic energy, burning to rapid oxidation, and electricity to movement of electrons.

One might now expect Churchland to tell us what property it is to which sensations, or some particular sensation, e.g., pain, reduces, or how we are to "explain consciousness in neurobiological terms".<sup>18</sup> But she offers no such account; she says only that attempts in this direction are "worth trying".<sup>19</sup>

Our remarks about water, H<sub>2</sub>O, and the explanatory gap show, however, that there is no reason to think such a project is even worth trying. We certainly expect to be able to find out more and more about the neural causes of pains, and of other sensations. But, as we have seen, this causal or correlational investigation offers no help in explaining how the *property* of being a pain could be identical to the *property* of being a set of neural activations of some kind. The net force of Churchland's remarks is thus no more than an expression of a hope that some future development will change the present situation, and lead us to the ability to explain the property identity that materialism implies.

At this point, it is often suggested that science is known for making progress, often in surprising directions, and that we should, therefore, trust in the progress of science to eventually provide a genuine explanation of a kind we cannot now even imagine. But here, I think, we should take our cue from Hume's (1748) essay "Of a Particular Providence and Of a Future State". To summarize very briefly, this essay is about the argument from design. A watch found on a beach would lead us to suppose an unseen watchmaker; how much more, then, should the exquisite organization of nature lead us to suppose an unseen designer of the world. Hume proceeds by conceding the point (at least for a while); but he insists that we are in no way entitled to infer any more perfection in the designer than what we observe in the design. If, therefore, we observe injustice and death in the actual world, the argument from design will fail to license an inference to an afterlife in which perfect justice is achieved. Matching the measure of the designer to the observed measure of perfection in the creation can lead us at most to suppose there is a designer that is subject to imperfections or limitations.

The point of this parable is to persuade you that we are not entitled to presume on the future of science beyond what we have observed in its actual development. And what we actually observe is that, while the term "explanatory gap" is of recent vintage, the problem it labels has been very well understood for over 300 years. Examples of the kinds of reductions Churchland points to have been available for over 200 years. Yet no progress at all has been made in understanding *how* it *could* be the case that sensations are nothing but neural events. Moreover, we know that science makes major advances in unexpected

---

<sup>18</sup> Ibid., p. 118.

<sup>19</sup> Ibid., p. 118.

directions. Physicists of 1850 could hardly have imagined what quantum mechanics or relativity theory would lead their successors to believe. So, even if we appeal to the idea that we will eventually make progress in science, we still have no reason to think that that progress will be in the direction Churchland, or any living materialist, imagines. It is *equally compatible* with belief in the progress of science to suppose that science will find a way to incorporate anti-materialism (denial of (M)) within an explanatorily satisfying worldview.

## V

There are many forms of materialism that we have not considered, and cannot consider in a single lecture. Thus, it is a large promissory note that I issue when I say that the essential points we have made, the centrality of (M), its inexplicability, along with its irrefutability, all hold up when variants of materialism are examined. What I want to do in closing is to consider what we should think if we agree in supposing that this promissory note can indeed be redeemed.

One might think that, in view of the points I have mentioned, materialism would be regarded with either agnosticism or doubt. As I have portrayed the matter, all that materialism has going in its favor is the fact that anti-materialists cannot *disprove* (M). So, one might be excused for suspending judgment about (M). Or, one might reasonably take the following attitude. The AM-argument is a strong, albeit not demonstrative, consideration against (M), and we cannot even understand how (M) could be true; thus, while agreeing that the falsity of (M) is not certain, we should think that (M) is probably false, and should try to find an overall worldview that incorporates both our sciences and the falsity of (M).

In actual fact, however, materialism is affirmed by the great majority of the best-known philosophers of our time. It is affirmed with passion and without qualification by any suspicion of doubt. The philosophers who affirm it are not only well known, they are, in fact, both knowledgeable and acute. They have expended a prodigious amount of genuine intellectual inventiveness and hard work in defense of the materialist point of view.

For all that, however, it seems clear to me that the degree of attachment to materialism is far in excess of what is warranted by the argumentation that is offered in support of it. The fact that the more defensible materialisms are the emptier materialisms goes unappreciated, and the attachment to the view remains strong even in those who believe that human beings will never understand how materialism could be true.<sup>20</sup> In short, in today's philosophical milieu, materialism has the status of an ideology—a view whose acceptance has become disconnected from the force of the reasoning that can be adduced in its favor.

I indicated above that I found this conclusion disturbing. The reason is, I hope, obvious. It cannot be any comfort to think that the best philosophers of one's day are in the grip of an ideology. The suspicion that this is so leads to the unpleasant thought that ideology is well-nigh unavoidable, at least in philosophy.

<sup>20</sup> See C. McGinn, *The Problem of Consciousness*.

Mere doggedness is not a virtue; but, perhaps, its avoidance is not a virtue to which we can honestly aspire. From the point of view of critical thinking, we must ask whether there is always a limitation, that is, whether there are always fundamental views that resist being brought under the scrutiny of critical thought. Since I am an optimistic person, I harbor the hope that these pessimistic suggestions are overdrawn. But I do think that the place of ideology in philosophical enterprises (especially those involving many contributors) and the question of avoidability of ideology are matters that deserve our serious consideration.

### Bibliography

Chisholm, R.

1957 *Perceiving: A Philosophical Study*, Cornell University Press.

Churchland, P. S.

1998 Brainishy: nonneural theories of conscious experience, in *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*, S. R. Hameroff, A. W. Kaszniak, and A. C. Scott, eds., MIT Press/Bradford.

Descartes, R.

1641 *Meditations on First Philosophy*. Translations taken from E. S. Haldane and G. R. T. Ross, *The Philosophical Works of Descartes*, Cambridge University Press, 1911.

Hume, D.

1748 *An Inquiry Into Human Understanding*.

Levine, J.

1983 Materialism and qualia: the explanatory gap, *Pacific Philosophical Quarterly* vol. 64, pp. 354-361.

Lycan, W. G.

1986 *Consciousness*, MIT Press/Bradford.

McGinn, C.

1991 *The Problem of Consciousness*, Blackwell.

Quine, W. V. O.

1960 *Word and Object*, Technology Press of MIT.

Robinson, W. S.

forthcoming *Understanding Phenomenal Consciousness*.

Velmans, M.

2000 *Understanding Consciousness*, Routledge.