

# Perspectival Representation and the Knowledge Argument

William G. Lycan

...this man wanted to be outside everything; to see everything hung in a vacuum, simply its own dead self.

--G.K. Chesterton, "The Shadow of the Shark"

Someday there will be no more articles written about the "Knowledge Argument" (Nagel (1974), Jackson (1982)). That is beyond dispute. What is less certain is, how much sooner that day will come than the heat death of the universe.

I thought I had said my own last words on the topic (Lycan (1987, Ch. 7), (1990b), (1996, Ch. 3)), but it seems not. There is at least a bit of unfinished business.

1. The irrepressible Mary. Let us quickly review Jackson's version of the argument. He offers the now painfully familiar example of Mary, the brilliant color scientist trapped in an entirely black-and-white laboratory. (Even she herself would have to be painted black and white.) There are potential complications that would have been avoided had Jackson made Mary entirely color-blind from birth instead of merely confined to her lab space; ink has been spilt over ways of getting Mary to experience color other than by seeing colored objects, which is an interesting topic but irrelevant to the point of the Knowledge Argument. So it is the color-blindness version I shall consider here. We simply stipulate that Mary is impaired in such a way that she cannot have color experiences of any sort.

Working through her modem and various black-and-white monitors, she becomes scientifically omniscient as regards the physics and chemistry of color, the neurophysiology of color vision, and every other conceivably relevant scientific fact; we may even suppose that she becomes scientifically omniscient, period. Yet when she is finally cured of her color-blindness, she sees colors themselves for the first time; and she thereby learns something, viz., she learns what it is like to see red and many of the other colors. That is, she learns what it is like to experience subjective or *phenomenal* redness, never mind the actual colors of the physical objects she encounters (which she already knew by using her spectrometer).

And so she has acquired information that is--by hypothesis--outside the whole domain of science. It is intrinsically subjective phenomenal information, and what she has learned seems to be an intrinsically perspectival *fact*, that eludes all of science. According to materialist theories of mind, no fact about the mind is anything but a neurophysiological or otherwise scientific or "objective" fact; so it would follow that materialism is false.

Actually there is a crucial distinction to be respected here. In Jackson's official formulation, the word "fact" is not used; his conclusion is just that "one can have all the physical information without having all the information there is to have" (p. 472). This, he says, refutes the doctrine he calls "Physicalism" (p. 469): "All (correct) information is physical information." But depending on how one construes the slippery term "information," "Physicalism" in this sense need not be taken to be an ontological claim at all (Horgan (1984)). It is most naturally understood as being about *truths*, rather than about what kinds of things there are; Flanagan (1992, p. 98) calls it "linguistic physicalism." And taken in that way, it is not entailed by materialism about the mind. (Materialism says only that human beings are made entirely of physical matter and that their properties, and facts about them, consist in arrangements of that

matter. It hardly follows that every sentence or proposition about a human being *means* something about physical matter.) However—witness his title—Jackson goes on to draw an explicitly ontological conclusion. I shall return to this distinction.

Here is a more formal statement of the Knowledge Argument, construed ontologically, as an objection to materialism.

(1) Before her cure, Mary knows all the scientific and other “objective” facts there are to know about color and color vision and color experience, and every other relevant fact. [Stipulation.]

(2) Upon being cured, Mary learns something, viz., she learns what it’s like (w.i.l.) to experience visual redness. [Seems obvious.]

\ (3) There is a fact, the fact of w.i.l. to experience visual redness, that Mary knows after her cure but did not know prior to it. [1,2]

(4) For any facts: if F1 = F2, then anyone who knows F1 knows F2. [Suppressed; assumes simple factive grammar of “know.”]

\ (5) There is a fact, that of w.i.l., that is distinct from every relevant scientific/“objective” fact. [1,3,4]

(6) If materialism is true, then every fact about color experience is identical with some physiological, functional, or otherwise scientific/“objective” fact.

---

\ (7) Materialism is not true. [5,6]

(4) is supplied because without (4), there seems no way to get (5) from (1) and (3).

2. Perspectivalism. There are three main materialist responses to the Knowledge Argument. One is brutally to deny (2) (Churchland (1985), Dennett (1991), Akins (1993)), and to insist that if Mary really did know *all* the scientific/“objective” facts, she would after all know w.i.l. to see red.

The idea is that despite her inability to see or even visually imagine colors, the scientifically omniscient Mary damn well could work out what it would be like to see them. At least, Jackson has given us no argument for thinking she could not. (Dennett reminds us that fantastical science-fiction examples of this kind are dangerous because we are taken in by the immediate image and fail to work out the real implications of the fantastical hypothesis.)

Jacksonian response: Yeah, OK, maybe. But it sure does seem that Mary would not know what it's like to see colors without having experienced color in some way.

A second rebuttal of Jackson is to grant (2) but balk at (3), holding that Mary’s acquisition is, not a fact, but a mere ability, a knowing-*how* (Nemirow (1990), Lewis (1990)) or a mere acquaintance (Conee (1994)).

The Ability theory has it that although Mary would acquire some knowledge, it would be only a knowing how, a skill or ability, not knowledge *that* anything, not knowledge of a fact. Jackson has done nothing to show that she has learned more than an ability to imagine colors, an ability to sort objects by color without using her spectrometer, etc.

Jacksonian response: Lycan (1996, Ch. 5) has given a number of arguments in favor of Jackson's claim that what Mary has learned is propositional and factual. One is that imagining is correct or incorrect. If Mary can imagine seeing a red object, this must mean imagining it correctly, not getting it wrong by imagining seeing what is in fact a different color. And presumably her ability is to be reliably correct, not just accidentally so. The best explanation of

that reliability is that she knows that that is how red objects look.

A second argument is linguistic: "Knowing wh-" locutions are true in virtue of the truth of "knowing that" locutions with referring terms in them. E.g., to know who robbed the diaper service is to know that N robbed the diaper service, for some suitable name N; to know when the pub closes is to know that the pub closes at t, where t refers to a time. So too, presumably, to know what it's like to see red is to know that it is like X to see red, for some suitable term X. Of course, it's hard, maybe impossible, to describe what it's like in English (except comparatively), but a demonstrative is natural here: If you're like me, you'll want to say, "It's like... THIS. I know what it's like; I just can't put it into words."

The third response is a generic that is by now standard; I shall call it the "perspectivalist" response. In one form or another, it has been suggested and/or defended by McGinn (1983), Horgan (1984), McMullen (1985), Churchland (1985), Van Gulick (1985), Tye (1986), Lycan (1987, 1990b, 1996), Loar (1990), Rey (1991), Levine (1998), and no doubt others. It is this third position that I shall continue to develop here.<sup>1</sup>

The perspectivalist begins by rejecting premise (4), the suppressed principle according to which (to put it another way) if someone knows that P but does not know that Q, then the fact that P and the fact that Q are different facts.

That principle may at first seem obvious. It seems to be licensed by Leibniz' Law: If fact F1 is known to Smith, and  $F1 = F2$ , then surely F2 is known to Smith. But there are clear counterexamples to it: The fact of water splashing just is the moving of lots of H<sub>2</sub>O molecules, but someone can know that water is splashing without knowing anything about H<sub>2</sub>O; the fact of my being overpaid just is the fact of WGL's being overpaid, but someone (who does not know that I am WGL) can know that WGL is overpaid while having no idea whether I am overpaid.

What has gone wrong? As always and notoriously, Leibniz' Law fails for *representation-dependent* properties. That Oedipus wanted to marry Jocasta but did not want to marry his mother does not show that Jocasta was not his mother; the poor woman was his marriage-object under one description or mode of presentation but not under the other.<sup>2</sup> And *being known to Smith* is a representation-dependent property: Whether Smith knows a given fact depends on how Smith represents that fact. She may know it under one representation but not know it under a different one. That is just what is going on in the "water" and "overpaid" examples. One may see water splashing but lack the chemical perspective entirely; less commonly, a mad chemist might record a motion of H<sub>2</sub>O molecules but be mentally so far removed from the perspective of everyday things and substances that she has no thought of water.

The "overpaid" example is perspectival too, but a different kind of perspective is at work. Someone can know that WGL is overpaid without knowing that I am overpaid, if that person has only a public, (non-auto-)biographical perspective on me and is not in a position to refer to me more directly. Even if the person were to come into the room, point straight at me and exclaim "*You* are overpaid," I might insist that the knowledge she thereby expresses is still not quite the same knowledge I have when I know that *I myself* am overpaid. (Especially if I believe that she is mistaken as to who I am.) As Hector Castañeda (1966) emphasized so many years ago,<sup>3</sup> if I myself am amnesic I may know many facts about WGL, including that he is overpaid, without knowing that I myself am overpaid; so it seems that what I know when I do know that I myself am overpaid is a different fact from any of those I could know while amnesic, and an intrinsically perspectival fact. In order to designate the person it is supposed to designate, a mental pronoun can be tokened only from a certain point of view; only I, WGL, can think "I" and thereby designate WGL.

Clearly, *being known to Mary* is a representation-dependent property; whether Mary knows a given fact depends on how she represents that fact. Facts can be differently represented from differing perspectives, and that is why (4) is false. Without (4), seemingly, the Knowledge Argument collapses.

But that is not the end of the story.

3. Fine-grained “facts.” Now, sometimes philosophers do individuate “facts” in a less chunky, more fine-grained sort of way. The late Roderick Chisholm (1976) used a version of principle (4) itself as a test for sameness of “fact.”<sup>4</sup> In his sense, the water splashing is a different fact from the H<sub>2</sub>O molecules’ moving, odd as it sounds to say that.

I do not think anyone could credibly insist that one of these two ways of counting “facts”—the chunky or the Chisholmian—is *correct* to the exclusion of the other. I make the following terminological proposal. Let us continue to use the term “fact” in the more usual chunky way, with which we began, and let us call Chisholm-facts “pieces of information.” The latter seems reasonable because Chisholm thinks of the things epistemically, as objects of conceptual knowledge rather than as chunks of the world. Thus, the fact of water splashing just is the moving of lots of H<sub>2</sub>O molecules, but that water is splashing and that H<sub>2</sub>O molecules are moving are two different pieces of information. Likewise, according to the materialist, that such-and-such neural goings-on are taking place in a subject’s brain and that it is like so-and-so for the subject to have an experience are the same fact but different pieces of information.

Once this distinction is introduced, the perspectivalist objection to the Knowledge Argument seems to have been blunted, for (by definition) (4) remains true at least of “facts” in the fine-grained sense, i.e., of pieces of information. Have Nagel and Jackson not then proved that there is *information* that is not public, objective, scientific etc. information?

I say they have: Just plug in “piece of information” for “fact” in our official statement of Jackson’s argument, and I believe, *pace* those who simply deny (2) and those who resist the inference from (2) to (3), that the resulting argument succeeds.

The existence of nonphysical, subjective, intrinsically perspectival etc. *information* may or may not be of metaphysical interest. But for purposes of philosophy of mind, here are two key ways in which it is not interesting: (i) The phenomenon is not specifically about the mind; it is everywhere. No amount of chemistry contains the information that *water* is splashing. No amount of economic etc. information contains the information that I am overpaid. And so forth. And/but (ii) it does not follow in any of these cases that the object or stuff in question—water, or WGL--has a nonphysical or immaterial *property*. I believe Jackson tacitly makes this inference, from nonphysical (piece of) information to nonphysical property, and so commits a fallacy; otherwise, he has no argument for the existence of epiphenomenal qualia.

4. The continuing problem. The perspectivalist has rebutted the Knowledge Argument by showing that its fourth premise is false of facts in the chunky sense. But, even once we have conceded and denigrated the point about nonphysical pieces of information, that is not enough. For we have not yet *shown how* the case of Mary in particular can be seen as an instance of the foregoing kind of Leibniz’-Law failure—i.e., as a case of knowing a fact under one representation but not knowing it under a different one. Also, as Charles Hill has acutely pointed out to me,<sup>5</sup> we have so far left it open for Nagel or Jackson to repair the argument by *qualifying* (4).

These two reasons are closely related, both being grounded in the fact that Mary’s

predicament does not seem akin either to the “water” example or to the “overpaid” example.

To expand the second reason: Though the “water” example is indeed a counterexample to (4), it is irrelevant, because it is a case of scientific ignorance, and we have stipulated that Mary suffers from no relevant scientific ignorance at all. So instead of (4) in full generality, Jackson could substitute,

(4') If  $F1 = F2$ , then anyone who knows  $F1$  *and is not suffering from scientific ignorance* knows  $F2$ .

Premise (2) already guarantees that Mary is not suffering from scientific ignorance, so this revised argument would defy the counterexample.

The “overpaid” example seems irrelevant too, because Mary does not seem to differ from the rest of us in virtue of any inability to use pronouns. So we can restrict (4) still further:

(4'') If  $F1 = F2$ , then, *barring pronominal discrepancies*, anyone who knows  $F1$  and is not suffering from scientific ignorance knows  $F2$ .

Neither the “water” example nor the “overpaid” example is a counterexample to (4''). So if we substitute (4'') for (4) and merely add the premise (4+) “Mary’s deficit is not a pronominal discrepancy,” Jackson’s argument goes through once more.

5. The perspectivalist approach to the continuing problem. To apply the standard two-perspectives explanation of Leibniz’-Law failure, we need a second perspective. That is, we already have Mary’s scientific perspective on everything physical, neurophysiological, functional etc. about seeing colors—all her various public, objective modes of representing the relevant facts. So what other perspective, what other mode of representation, could explain her failing to know w.i.l.?

My answer: An introspective perspective.<sup>6</sup> As Nagel emphasizes, to know w.i.l., one must either have had the experience oneself, in the first person, from the inside, or been told w.i.l. by someone who has had it and is psychologically very similar to oneself. The descendants of Descartes of course champion that first-person perspective and believe it is the one we are stuck in whenever we philosophize about the mind. But what can a materialist say about any first-person perspective? Materialists from Ryle onward have seemed almost to deny the existence of such a thing.<sup>7</sup>

How might a materialist reverse that attitude, and implement an introspective perspective? Why, in good neural hardware. That is where theories of conscious awareness and introspection come in.

Occurrent mental or psychological states fall roughly into four categories: those whose subjects are consciously attending to them; those whose subjects are aware of being in them though not actively introspecting them; those whose subjects are not aware of being in them, but could have been had they taken notice; and those, such as language-processing states, which are entirely subterranean and inaccessible to introspection. A theory is needed to explain these differences.

D.M. Armstrong’s (1968, 1981) Lockean “inner sense” or “higher-order perception” theory fills the bill admirably. (See also Lycan (1987, Ch. 6) and (1996, Ch. 2), as well as the “higher-order thought” theory defended by Rosenthal (1993)). According to such theories, a subject’s awareness of her/his own mental state consists in a representing of that state itself. The “inner sense” view has it that the representing is done quasi-perceptually, by a set of functionally

specified internal attention mechanisms of some kind that scan or monitor “first-order” mental/brain states. This theory explains the differences between our four categories of first-order state in the obvious way.

(If you are inclined to doubt whether awareness or introspection of a mental state really requires higher-order quasi-perception or at least representation, consider: Awareness is intentional; it is always awareness of something. In particular, awareness of a mental state one is in has that mental state as its intentional object. If you doubt that such awareness consists in higher-order representation, i.e., in some representation of the first-order state in question, you are abandoning the thesis that intentionality is representation. In my view that is not an easy thesis to abandon. Descartes himself championed it.)

In addition to that quick & clean argument, and the fact that Higher-Order-Representation theories do explain the differences between our four types of mental state, there are other features that recommend HOR theories. I have expounded them in previous works (chiefly Lycan (1996: 15ff.)), so I shall only mention a couple of them here. For example: such theories explain a nagging ambiguity and/or dialect difference regarding sensations and feeling; and they afford what I believe is the best going account of “knowing what it’s like” to experience a particular sort of sensation (Lycan (1996; Chapter 3).

It is worth heading off two bad and ignorant objections. These two have been scouted more than once before, but I have found that they still come up every time a HOR theory is propounded. First is the Qualia-Creation objection: “You say that what makes a state conscious is that the state is the intentional object of a higher-order perception or thought. But surely the mere addition of a higher-order representation cannot bring a sensory or phenomenal quality into being. If the original state had no qualitative or phenomenal character, the HOR theories do not in any way explain the qualitative or phenomenal character of experience; a mere higher-order representation could hardly bring a phenomenal property into being. So how can they claim to be theories of *consciousness*?”

But it is no claim of either HOR theories per se to have explained anything about qualitative character; they are theories only of the distinction between mental states one is aware of being in and mental states one is not aware of being in. Some other theory must be given of phenomenal character. (Some critics have supposed that HOR theories have been aimed at other explananda, in particular at the sensory qualities; the critics have expressed understandable incredulity at the idea that a mere higher-order thought or even a higher-order quasi-perception could explain the original qualitative character of a sensory state. But I know of no HOR theorist who has claimed to explain anything but awareness in the sense specified above—certainly not Armstrong, myself, or Rosenthal.)

Second, the Regress objection. Some philosophers have feared a regress. If the second-order representation is to confer consciousness on the first-order state, it must itself be a conscious state; so there must be a third-order representation of it, and so on forever. But HOP and HOT theorists reject the opening conditional premise. The second-order representation need not itself be a conscious state. (Of course, it may be a conscious state, if there does happen to be a higher-order representation of it in turn.)

If there is something like an intra-mental sense modality, then it would come with an introspective mode of representation, and it would present its objects from a special, uniquely internal point of view. This introspective perspective would be unlike any public point of view available to Mary. The internal monitor would say things like “You’re having one of *those* again,” and “*This* is what it’s like to have one of those” (except of course that it would not use

English words, but would mobilize representations of its own proprietary kind).

On this view, the perspectival difference would after all be much like the pronominal discrepancies. “You can know all the scientific stuff, but no amount of that will tell you what *this* is like” (said from the inside) sounds much like “You can know all the economic, biographical facts about WGL, but no amount of that information will tell you anything about *me*” (said by an amnesic who is, in fact, WGL), and like “You can know all about Caldwell Hall and all its rooms and offices and know that the meeting will be held in room 213, but no amount of that information will tell you that the meeting will be held *right here*.”

It is not just a matter of pronouns, of course, even if one of them is a mental pronoun. Introspection type-classifies mental states—that is, it classifies the various states it surveys into states of the same or different kinds—in its own distinctive way, under its own representations of them (“There’s one of those *semanthas* again”). But those representations are not going to be synonymous with English words, because your internal monitors don’t speak English or any other natural language. (The reason, I have argued elsewhere, is that no English sentence can have both the same chunky fact as its referent and the same mode of presentation, where “modes of presentation” are construed psychologically, as inferential or otherwise functional roles. This is analogous to the intrinsic perspectivalness we see in the pronominal cases: A token of the expression “my left foot” has both a distinctive inferential role and a particular referent that nowhere else coincide.) Thus, the introspective representations are not going to be deducible from neuroscience or from any other body of public information expressed in English; that is why I believe Jackson is right about nonphysical pieces of information.

This perspectivalist picture also explains two things that badly need explaining: First, the seeming ineffability of w.i.l. One often cannot express in words what it is like to have a particular sensation. What is it like to experience the yellowy-orange of a yellowy-orange after-image? “It’s yellowy-orange, yes, I heard you the first time”—but can you tell someone in intrinsic, not comparative, terms, what it’s like to experience visual yellowy-orange? If you are like me, at this point you dummy up: “It’s like...*this*. I know what it’s like; I just can’t put it into words.”

The second explanandum is Levine’s (1983, 1993) “explanatory gap.” As you can tell, I believe he is right to contend that there is one. One can’t explain what is ineffable. No science is or could be formulated in Introspectorese. Notice that there are parallel explanatory gaps in the pronominal cases too: If I am amnesic, you can explain till you’re blue in the face why William G. Lycan is in New Zealand but, so long as I am unaware that I am WGL, you will not have explained the truth I enunciate when I observe that *I* am in New Zealand. Notice too that no extravagant metaphysical conclusions can be drawn from any such gaps. William G. Lycan and I are simply, numerically identical.

So, I’ve granted, the introspective representations’ contents are going to be nonphysical pieces of information. And that, says the perspectivalist, is why Mary can know all the scientific stuff about experiences of red and not know w.i.l. to experience red: It is because, never having had a first-order experience of red herself, she cannot represent that experience from the introspective perspective. She can represent other people’s experiences of red, but only from the neuroscientific, the commonsensical, and other public perspectives. Once she is released and does experience red herself, she then does introspect and represent that experience in the distinctive first-person way—and that is when and how she comes to know w.i.l. to experience red, gaining a nonphysical piece of information.

6. And (4'')? If the difference between Mary Before and Mary After is not just pronouns, then it is not excluded under the “pronominal discrepancy” clause. If (4'') is true, then, the argument would still go through. So the perspectivalist must argue that even (4'') is too strong. And s/he cannot do that by just insisting that (4'') is falsified by the case of Mary, because that would beg the question.

Let us talk of a “quasi-perceptual” perspective. Sense organs and other detectors, monitors etc. are intrinsically perspectival and deliver only intrinsically perspectival representations, in that (a) they show the objects they detect only from particular points of view, (b) they key on different properties of the objects, (c) they give you different packets of information about the same objects, and (d) they offer their representational contents under different modes of presentation. Suppose you are in the forest at night and, very close by, a tree falls and crushes your Porsche.<sup>9</sup> The falling of a tree is, scientifically speaking, the complex motion of gazillions of cellulose etc. molecules (same fact). But that event can be thought of and described nonscientifically, under many different sensory modes of presentation: If you have good night vision, you see the tree, falling. What you hear is a great rending crash. You feel the impact of the shock wave in the air. (You can also think of the event as “what destroyed my car.”) These are all different pieces of information, but each describes the same chunky event or fact, scientifically just the motion of molecules.

This refutes (4''), I believe. Without either scientific ignorance or pronominal discrepancy, someone could know that the tree had fallen without knowing that there had been an audible crash, etc., etc.

Armstrong’s “inner sense” is not exactly like an external sense such as sight, hearing or touch. (The differences stem from the fact that its function is not to detect adaptively significant features of objects in the external environment.) But it is analogous; it is a *quasi*-perceptual faculty, a form of proprioception. And part of the analogy is that it gives a perceptual-type perspective on and mode of presentation of a brain fact, different ones from those of science itself, sight, hearing, touch and so on. So if we were to restrict 4'' further, to exclude differences of sensory perspective, I would argue that quasi-sensory perspective should count as well.

Thus, were Jackson to replace (4'') by

(4''') If  $F1 = F2$ , then, barring pronominal discrepancies and differences of sensory or quasi-sensory perspective, anyone who knows  $F1$  and is not suffering from scientific ignorance knows  $F2$ .

he would have to add the premise (4++) “Mary’s deficit is not the lack of some sensory or quasi-sensory perspective”; and this time his argument would not go through, because (4++) is not granted by the perspectivalist, and should not readily be granted by anyone.

7. But it's not over. I have attacked the revised Jacksonian principle (4'') by offering the falling-tree counterexample, to illustrate the case of differing sensory perspectives on the same physical event. But it may be objected<sup>10</sup> that if a person *is not suffering from scientific ignorance*, i.e. is scientifically omniscient, and does know that one of the relevant aspects (the visual, a dark, heavy motion of something massive, the aural, a rending crash, the tactual, a shock wave, etc.) occurred, then, through scientific omniscience, that person will know that each of the other things occurred also, so my case does not after all yield a counterexample to (4'').

One issue needs to be disposed of before I can address that objection directly: There is an ambiguity in the idea of “scientific omniscience.” I have been thinking of it as omniscience



regarding scientific *theory and principles*, of the kind we get from science textbooks. Scientific omniscience in that sense would not necessarily allow our subject to infer anything about a tree falling, from a shock wave or from a great crash, because nothing in physics or any other science could tell her what particular worldly event caused the shock wave or the crash; the latter is specific information about a one-time historical occurrence; so the counterexample stands. But let us waive this, because for his purposes Jackson could have built into his example that Mary knows all *particular* “objective” facts as well as scientific theory and principles. Let us also stipulate that logic is included in scientific omniscience.

Now, there is a simple answer to the objection: The foregoing formulation of (4'') was a bit loose. (4'') was not intended as denying that a scientifically omniscient person could *work out* other descriptions from the one, by scientific reasoning. Like the original (4), (4'') was motivated by Leibniz' Law, though qualified against two types of representation-dependent properties (those illustrated by the “water/H<sub>2</sub>O” and “overpaid” examples); the Knowledge Argument infers “There exists a non-scientific/‘objective’ fact” from “All the scientific/‘objective’ facts were already known to Mary” and “Mary has learned a new fact.” So of course (4'') should have been formulated as, “If F1 = F2, then, barring pronominal discrepancies, anyone who knows F1 and is not suffering from scientific ignorance *thereby* (eo ipso or ipso facto) knows F2”; *working out* F2 does not count.

However, that simple answer invites (indeed screams for) a further Jacksonian weakening of (4'''). Here is (4''''): “If F1 = F2, then, barring pronominal discrepancies, anyone who knows F1 and is not suffering from scientific ignorance *can work out* and thereby *come to know* F2.” For the foregoing reason, (4''') resists the falling-tree example, but (4''''') also preserves the Knowledge Argument.

Moreover, (4''''') preserves the Argument in a more damaging way than usual. I cannot just apply the usual formula and fashion a counterexample based on a new sort of mode of presentation, exploiting the representation-dependence of “knows” as usual. (I had been confident that since the original argument was still basically a Leibniz'-Law argument and *being known to Mary* is a representation-dependent property, there would always be a fix.) Because, notice, the weakening in (4''''') is not, like the previous weakenings in (4') and (4''), an added condition on the F1-knowing subject. Rather, it itself takes the form of weakening the consequent inside the universal generalization. So there is no obvious counterexample based on a new sort of mode of presentation.

By the same token, (4''''') is not at all motivated by Leibniz' Law or by any considerations of sameness/difference of fact. The focus has shifted. The new Jacksonian problem posed by (4''''') for the materialist is that, so far as has been shown, facts of w.i.l. cannot be worked out even by a scientific omniscient. And this problem is familiar in its own right: It is the failure of *logically* reductive explanation, as touted by Chalmers (1996). It is also closely akin to the Explanatory Gap.

I agree (contra Churchland, Dennett and Akins) that facts of w.i.l. cannot be worked out from any body of scientific/“objective” information, so I say that if (4''''') is true, the newly revised Knowledge Argument succeeds. But I do have a fix, an objection to (4'''''); and it is not ad hoc, but is based on views I have already defended in Lycan (1996). The reason facts of w.i.l. cannot be worked out from any body of scientific/“objective” information is that they are ineffable, and I gave a reason independent of the present issue why they are so; yet (I also argued) their ineffability and the reason for it are entirely compatible with materialism. Therefore, so far as has been shown, (4''''') is falsified by the putative a posteriori identity of its being like

[whatever] to experience visual redness with the relevant complex functional etc. fact. The Knowledge arguer could now go to (4'''''): "If  $F1 = F2$ , then, barring pronominal discrepancies *and so long as  $F2$  is effable*, anyone who knows  $F1$  and is not suffering from scientific ignorance can work out and thereby come to know  $F2$ ." Quite possibly, but the arguer would need the further premise that the fact of its being like [whatever] to experience visual redness is effable, which I reject.

So materialism remains unrefuted by this latest hedged version of the Knowledge Argument. But we have given some aid and comfort to the enemy, for two reasons. First, I have had to appeal to theory to impugn (4'''''). (4), (4') (4'') and (4''') were refuted by plain counterexamples of the type familiar from the Leibniz'-Law literature, but the worst I can say of (4''''') is that it is incompatible with a plausible view that I have also defended elsewhere. Second, I have just about vindicated an antimaterialist idea defended by Chalmers (1996): that although there are a number of counterexamples to (4) and to (4') and to (4'') and even to (4'''), they work on fairly simple and superficial principles, while facts of w.i.l. are at best a deeper phenomenon, and are probably a unique counterexample to (4''''') *if* they are a counterexample at all.

On that second point, I am undaunted and unapologetic. *Each* type of counterexample to the (4)-family principles is a unique type. And Lycan (1996) not only admitted but insisted that consciousness and k.w.i.l. are very special phenomena; my own theory predicts that and explains why they are. The worst that can be said is that they are not simple and superficial counterexamples to (4)-(4'''''), but are of a deeper though still tractable sort.

8. A curious consequence. I said that before she was cured, Mary could represent other people's experiences of red, but only from the various public perspectives and not introspectively. But perhaps that is only a contingent limitation. Assume that Mary has a perfectly good internal monitor or set of them, which she uses to introspect the first-order mental states she does have. Then we can imagine a futuristic surgical procedure in which we wire up the relevant internal monitors of Mary's into some other person's visual cortex. If that is possible in principle, then on the "Inner Sense" view, Mary would be able to introspect the other person's experiences of red, and find out in that way w.i.l. see red.

Armstrong himself observed (in Armstrong & Malcolm 1984, pp. 113-14) that the Inner Sense view makes the privacy of experience a contingent fact. He went on to speculate as to what it would be like to "introspect," or extrospect, someone else's first-order sensations, such as pains. There are two possible subcases. First, it might be like direct awareness of an ordinary pain of your own, accompanied by the same reactions you have to your own pains. The other person's pain would in effect become your own pain, as if the other's ailing body part were an extension of your own body.

Second, the intro- or extrospective awareness might only generate a comparatively abstract sympathy; you would still be more of an onlooker than a sufferer. You would detect the quale or qualitative sensory core of the other person's pain, but it would not have the effects in you that your pains normally have—no intrinsic desire that it cease, no distraction, no impulse to bodily pain-behavior. Such dissociation between the core qualitative character of pain and the affective components of pain actually occurs within individual human subjects, e.g., those receiving morphine; they reliably report that the pain is still there and has not diminished in intensity, but that they no longer mind it and can take an entirely detached attitude toward it. (In her recent book on pain, Valerie Gray Hardcastle argues that there is a good neurophysiological reason for

this sort of split.)

As between the two possibilities, I believe the first is very unlikely. My reason is functionalist: *Presenting itself to introspection* and *causing the belief that one is in pain* are two elements of a pain's functional role, but only two. Other elements on the output side include, as before, causing the desire that it cease, distracting the attention, and causing wincing, favoring etc. behavior. The latter effects would be produced in the suffering subject you are intro- or extrospecting, but there is no reason to think they would be produced in you, for in normal subjects they do not depend on the pain's being introspected or known of, but occur even when one is unaware of having the pain. Therefore, as between Armstrong's two scenarios, I think the second is by far the more likely.

The possibility that privacy could be overcome by futuristic surgery seems plausible to me, but then I've been defending Armstrong's view for a long time. If it does not seem plausible to you, then you may find that you have an objection either to the "Inner Sense" view itself or to its incorporation into the perspectivalist theory of knowing w.i.l.

And of course, any objection to the "Inner Sense" view is an objection to the perspectivalist theory as formulated here. It is a further question whether the perspectivalist could swap "Inner Sense" for a different theory of conscious awareness, such as Rosenthal's, and still be able to make the perspectivalist move. I believe that would work, since I hold that all awareness is intentional and that all intentional states represent their intentional objects under distinctive modes of presentation; but I shall not try to make that case here.

### Footnotes

1 For a battery of arguments against the Nemirow-Lewis "ability" theory, see Lycan (1996, Ch. 5).

2 I here ignore two complications: that some theorists insist that a "representation-dependent property" is no genuine property at all, and that Leibniz' Law does hold for properties incorporating representations when the representations are de re.

3 Following Geach (1957); see also Perry (1979).

4 Chisholm himself preferred to speak of *states of affairs*; for him, a fact is a state of affairs that (actually) obtains.

5 Hill's observation directly inspired this paper; I thank him again for it.

6 N.b., rhymes with "insective detective."

7 As is vociferously deplored by Searle (1992); see also Siewert (1998).

8 There are parallel explanatory gaps in the pronominal cases too. Notice that no extravagant metaphysical conclusions can be drawn from any of them, for much the same reason, that nonphysical pieces of information are metaphysically harmless: The gaps are due to there being

intrinsically perspectival, nonphysical pieces of information, which itself does not show that anything has a nonphysical or immaterial property. (However, for an argument that there may be further metaphysical trouble, see Levine (in press).)

I do not here deal with the far more elaborate anti-materialist argument based on the explanatory gap, due to Jackson (1993) and Chalmers (1996). See Lycan (in press), as well as Block and Stalnaker (1999) and Levine (in press).

9 Yes, I know that in real life you do not have a Porsche, or you would not be reading this article. Or, if you do actually own a Porsche, would you be willing to send me a check for \$1,000?

10 As it was by Amy Kind, at the NEH Summer Institute “Consciousness and Intentionality,” University of California, Santa Cruz, June, 2002.

### References

- Akins, K. (1993). “What Is it Like to be Boring and Myopic?” In B. Dahlbom (ed.), *Dennett and His Critics*. Oxford: Basil Blackwell.
- Armstrong, D.M. (1968). *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul.
- Armstrong, D.M. (1981). “What is Consciousness?” In *The Nature of Mind*. Ithaca, NY: Cornell University Press.
- Block, N., and R. Stalnaker (1999). “Conceptual Analysis, Dualism, and the Explanatory Gap.” *Philosophical Review* 108: 1-46.
- Castañeda, H.-N. (1966). “‘He’: A Study in the Logic of Self-Consciousness.” *Ratio* 8: 130-57.
- Chalmers, D. (1996). *The Conscious Mind*. Oxford: Oxford University Press.
- Chisholm, R. (1976). *Person and Object*. London: George Allen and Unwin.
- Churchland, P.M. (1985). “Reduction, Qualia, and the Direct Introspection of Brain States.” *Journal of Philosophy* 82: 8-28.
- Conee, E. (1994). “Phenomenal Knowledge.” *Australasian Journal of Philosophy* 72: 136-50.
- Davies, M., and G. Humphreys (eds.) (1993). *Consciousness*. Oxford: Basil Blackwell.
- Dennett, D.C. (1991). *Consciousness Explained*. Boston, MA: Little, Brown & Co.
- Flanagan, O. (1992). *Consciousness Reconsidered*. Cambridge, MA: Bradford Books / MIT Press.
- Geach, P. (1957). “On Belief about Oneself.” *Analysis* 18: 23-24.
- Horgan, T. (1984). “Jackson on Physical Information and Qualia.” *Philosophical Quarterly* 34: 147-83.
- Jackson, F. (1982). “Epiphenomenal Qualia.” *Philosophical Quarterly* 32: 127-36. Reprinted in Lycan (1990a); page references are to the latter.
- Jackson (1993). “Armchair Metaphysics.” In M. Michael and J. O’Leary-Hawthorne (eds.), *Philosophy in Mind*. Dordrecht: Kluwer Academic Publishing.
- Levine, J. (1983). “Materialism and Qualia: The Explanatory Gap.” *Pacific Philosophical Quarterly* 64: 354-61.
- Levine, J. (1993). “On Leaving Out What It’s Like.” In Davies and Humphreys (1993).

- Levine, J. (1998). "Conceivability and the Metaphysics of Mind." *Noûs* 32: 449-80.
- Levine, J. (in press). *Purple Haze: The Puzzle of Conscious Experience*. Oxford: Oxford University Press.
- Lewis, D. (1990). "What Experience Teaches." In Lycan (1990a).
- Loar, B. (1990). "Phenomenal States." In Tomberlin (1990).
- Lycan, W.G. (1987). *Consciousness*. Cambridge, MA: Bradford Books / MIT Press.
- Lycan, W.G. (ed.) (1990a). *Mind and Cognition*. Oxford: Basil Blackwell.
- Lycan, W.G. (1990b). "What is the 'Subjectivity' of the Mental?" In Tomberlin (1990).
- Lycan, W.G. (1996). *Consciousness and Experience*. Cambridge, MA: Bradford Books / MIT Press.
- Lycan, W.G. (in press). "Vs. a New A Priorist Argument for Dualism," forthcoming in E. Sosa and E. Villanueva (eds.), *Philosophical Issues, Vol. 13* (Oxford: Basil Blackwell).
- McGinn, C. (1983). *The Subjective View*. Oxford: Oxford University Press.
- McMullen, C. (1985). "'Knowing What It's Like' and the Essential Indexical." *Philosophical Studies* 48: 211-34.
- Nagel, T. (1974). "What Is It Like to Be a Bat?" *Philosophical Review* 82: 435-56.
- Nemirow, L. (1990). "Physicalism and the Cognitive Role of Acquaintance." In Lycan (1990a).
- Perry, J. (1979). "The Problem of the Essential Indexical." *Noûs* 13: 3-21.
- Rey, G. (1991). "Sensations in a Language of Thought." In E. Villanueva (ed.), *Philosophical Issues, I: Consciousness*. Atascadero, CA: Ridgeview Publishing.
- Rosenthal, D. (1993). "Thinking that One Thinks." In Davies and Humphreys (1993).
- Searle, J. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- Siewert, C. (1998). *The Significance of Consciousness*. Princeton, NJ: Princeton University Press.
- Tomberlin, J.E. (ed.) (1990). *Philosophical Perspectives, Vol. 4: Action Theory and Philosophy of Mind*. Atascadero, CA: Ridgeview Publishing.
- Tye, M. (1986). "The Subjective Qualities of Experience." *Mind* 95: 1-17.
- Van Gulick, R. (1985). "Physicalism and the Subjectivity of the Mental," *Philosophical Topics* 13: 51-70.