## Forthcoming, Philosophical Issues 13, 2003:163-204

# A Naturalist-Phenomenal Realist Response To Block's Harder Problem Brian P. McLaughlin Rutgers University

Ned Block (2002) claims that there is "an epistemic tension" between two fairly widely held commitments: to phenomenal realism and to naturalism. Phenomenal realism is the view that (a) we are phenomenally consciousness, and that (b) there is no a priori or armchair sufficient condition for phenomenal consciousness that can be stated (non-circularly) in nonphenomenal terms (p.392). Block points out that while phenomenal realists reject "armchair philosophical reductive analyses" (p.393) of consciousness—such as analytical functionalism—"phenomenal realists have no brief against scientific reduction of consciousness" (p.393). His characterization of naturalism is complex:

Naturalism is the view that it is a default that consciousness has a scientific nature (and that similarities in consciousness have scientific natures). I shall assume that the relevant sciences include physics, chemistry, biology, computational theory, and parts of psychology that do not explicitly involve consciousness. (The point of the last condition is to avoid the trivialization of naturalism that would result if we allowed the scientific nature of consciousness to be...consciousness.) I shall lump these sciences together under the heading 'physical', thinking of naturalism as the view that it is a default that consciousness is physical (and that similarities in consciousness are physical). So naturalism = default physicalism, and is thus partly an epistemic thesis...My naturalist is not a 'die-hard' naturalist, but rather one who takes physicalism as a default, a default that can be challenged. My rationale for defining 'naturalism' in this way is that this version of the doctrine is plausible, widely held, and leads to the epistemic tension that I am expositing. (p.398)

Here and elsewhere in this paper, inserted pages references are to Block 2002.

What is it to be phenomenally conscious? What are states of phenomenal consciousness? Here is a standard characterization: "We can say that a being is...phenomenally conscious...when there is something it like to be that being. A mental state is [phenomenally] conscious when there is something it is like to be in that state. Conscious states include states of perceptual experience, bodily sensations, mental imagery, occurrent thought, and more...Each of these states has a phenomenal character, with phenomenal properties (or qualia) characterizing what it is like to be in the state "(Chalmers 2002, p.248). (Emphases his.) Block offers a useful clarification of one issue: "Imagine two persons both of whom are in pain, but only one of whom is introspecting his pain state and is, in that sense, conscious of it. One could say that only one of the two has a conscious pain. This is not the sense of 'conscious' used here. In the sense of 'conscious' used here, just in virtue of having pain, both have conscious states. To avoid verbal disputes, we could call the sense of 'conscious' used here *phenomenality*. Pains are intrinsically phenomenal and in that sense are intrinsically conscious. In that sense—but not in some other senses—there cannot be an unconscious pain." (p.394) (Emphases his.) By 'pain' I take Block to mean the felt quality of pain; and I take him to be maintaining that someone can have that quality, even when he or she is not introspecting it. I shall not employ Block's idiosyncratic term 'phenomenality', but when I speak of phenomenal consciousness, or simply of consciousness, I shall have in mind what he calls phenomenality. Thus, being in pain, for instance, suffices for being phenomenally conscious.

In saying that his "naturalist is not a 'die-hard' naturalist, but rather one who takes physicalism as a default," I take it that Block means that his naturalist is not a die-hard physicalist, but rather one who takes physicalism as a default. What, then, is meant by 'default'? He says: "A view on a given subject is the default if it is the only one for which background considerations give rational ground for tentative belief." (p.393) Thus, the naturalist holds that as concerns phenomenal consciousness, physicalism is the only view for which background considerations give rational ground for tentative belief. By 'a tentative belief', Block means a belief held by the believer in such a way that as to be subject to rational revision—a non-dogmatic belief.<sup>3</sup> In answer to the question, "What does 'grounds for rational belief' mean?"(p.407), Block answers: "I take this to be an epistemic level that is stronger than 'reasons for believing' and weaker than 'rational certainty'. I take it that a ground of rational belief that p allows knowledge that p but mere reason for believing p does not" (p.407). The issue of what rational grounds for belief are adequate for knowledge is, of course, deeply controversial. Block says nothing at all about this matter. To put the difficult issue of grounds for knowledge aside, then, naturalism, in Block's sense, is just the view that, as concerns phenomenal consciousness, physicalism, while open to rational challenge, is the only view for which background considerations provide rational grounds for belief.<sup>4</sup> Thus, as Block tells us at one point, his "naturalism is close to what John Perry (2001) calls 'antecedent physicalism'" (p.398). Naturalism in this sense is, I believe, fairly widely held, though I am uncertain whether it is the majority position. In any case, I am a naturalist-phenomenal realist. And Block seems to consider himself one as well. What follows is mainly addressed to other naturalist-phenomenal realists.

The epistemic tension between phenomenal realism and naturalism that Block claims to have found—a tension that I have yet to characterize—does not, he says, lead to paradox (p.424). He thus does not purport to show that phenomenal realism and naturalism are incompatible. Indeed a naturalist-phenomenal realist could, he adds, learn to "live with" the epistemic tension (p.424). The alternative, however, is to "reject or restrict the assumption of naturalism or of phenomenal realism" (p.424). And he states that given the epistemic tension, "phenomenal-realist naturalists may want to weaken their commitment to naturalism or to phenomenal realism" (p.393). My main concern in this paper will be with what, exactly, Block takes the epistemic tension in question to be, and whether it puts rational pressure on naturalist-phenomenal realists to compromise their commitment either to naturalism or to phenomenal realism.

Block's claim of an epistemic tension here is, at first blush, puzzling. There is, to be sure, a question whether consciousness has a physical nature (even in Block's very broad sense of 'physical'). So we might wonder whether we have rational grounds for belief in physicalism. But why think that naturalism (i.e., default physicalism) is in epistemic tension with phenomenal realism? There are concepts that lack noncircular a priori or armchair sufficient conditions of application yet are such that we take the default

Block confirmed this reading of 'tentative belief'; he has assured me that a tentative belief is a belief, not a doxastic state that is somehow weaker than belief (than outright belief).

<sup>&</sup>lt;sup>4</sup> Notice that a dogmatic physicalist—a "die-hard" physicalist—would not count as a naturalist in Block's sense. Nothing turns on this unhappy consequence, however.

to be that they pick out properties or kinds that have a scientific nature. The concept of water is an example. There is no noncircular a priori or armchair sufficient condition for being water, yet we take the default to be that water has a scientific nature; indeed we take ourselves to know what its scientific nature is.

By 'an a priori or armchair sufficient condition', however, Block means not only a priori (or armchair) metaphysically sufficient conditions, but also metaphysically contingent a priori sufficient conditions. Thus, the phenomenal realist holds that concepts of phenomenal consciousness lack even contingent a priori reference-fixing conditions. (The only a priori reference-fixer for the concept of pain, for instance, is the felt quality of pain itself.) It has been argued that the concept of water has a contingent, a priori reference-fixing condition, and that its having such a condition figures in our justification for believing that water =  $H_2O$ .<sup>5</sup> And it has been argued that the absence of either a priori metaphysically sufficient conditions for states of phenomenal consciousness or contingent a priori reference-fixing conditions for concepts of phenomenal consciousness, results in an *explanatory gap* between physical states and states of phenomenal consciousness.<sup>6</sup> The appearance of such an explanatory gap reveals an epistemic tension between naturalism and phenomenal realism. This epistemic tension, of course, has been a topic of intense discussion for many years.

While Block holds that there is indeed an appearance of an explanatory gap between physical states and states of phenomenal consciousness, he does not think that it results from the failure of concepts of phenomenal consciousness to have contingent a priori reference-fixing conditions. By his lights, the concept of water lacks any contingent a priori reference-fixing condition, yet we know that water =  $H_2O$ .

But in any case, while I shall return in part II to the explanatory gap, it will not be my primary concern in this paper. The reason is that Block claims to have identified an epistemic tension between phenomenal realism and naturalism that is different from the appearance of an explanatory gap (p.418). It is that epistemic tension that will be my primary concern.

Block claims that the epistemic tension in question poses a problem for the naturalist-phenomenal realist that is distinct from, and additional to, "the hard problem," a problem that he formulates as follows:

The hard problem is one of explaining why the neural basis of a phenomenal quality is the neural basis of *that* phenomenal quality rather than another phenomenal quality or no phenomenal quality at all (p.394) (emphasis his)

<sup>6</sup> See, e.g., Levine 1983, and 2000; and Chalmers 1995, 1996, 2002.

<sup>&</sup>lt;sup>5</sup> See Jackson 1993, and Chalmers 1996.

<sup>&</sup>lt;sup>7</sup> See Block and Stalnaker 1999. They deny, as well, that concepts such as the concept of life have either a priori metaphysically sufficient conditions of application or contingent a priori reference-fixing conditions of application (1999, p.000).

Here Block follows David Chalmers (1995, 1996). And I shall follow them both in calling this "the hard problem." Block thinks that, "we can see a glimmer of hope for a how a solution [to the hard problem] might one day be found" (p.394). As concerns the problem posed by the epistemic tension he claims to have identified, however, he thinks we have no conception of how it could be solved—or at least no conception short of compromising or weakening either naturalism or phenomenal realism. He thus calls his problem,

#### I. The Harder Problem.

By way of contrasting the harder problem with the hard problem, Block says:

The harder problem, as I shall call it, is more epistemological than the hard problem. A second difference: the hard problem could arise for someone who has no conception of another person, whereas the harder problem is tied closely to the problem of other minds. Finally, the harder problem reveals an epistemic tension or at least discomfort in our ordinary conception of consciousness which is not suggested by the hard problem, and so in one respect is harder. Perhaps the harder problem includes the hard problem and is best thought of as an epistemic add-on to it. Or perhaps they are in some way facets of a single problem. Then my point is that this single problem breaks into two parts, one of which is more epistemic, involves other minds, and involves an epistemic tension. (pp.391-2) (Emphasis Block's.)<sup>8</sup>

He takes "the harder problem" to differ from the problem of other minds in that "the harder problem," unlike that problem, is not a skeptical problem; but of this, more later.

The hard problem would arise should the neuro-scientific project of identifying neural bases for states of phenomenal consciousness in human beings and members of neighboring species succeed. Another problem that would remain is the problem of the nature of phenomenal consciousness. That is the problem that Galen Strawson (1996) had in mind when he spoke of the hard-part of the mind-body problem, the problem that Colin McGinn (1989) had in mind when he spoke of the hard-nut of the mind-body problem, and the problem that Thomas Nagel (1974) had in mind when he said that consciousness is what makes the mind-body problem hard. Block would not, I believe, claim to have identified an epistemic tension for naturalist-phenomenal realism that is "an epistemic add on" the problem of the nature of phenomenal consciousness. He would, I believe, acknowledge that the epistemic tension that he is concerned with arises when we try to determine the nature of phenomenal consciousness; and he would, I also believe, acknowledge that if we knew the nature of phenomenal consciousness, we could resolve

<sup>&</sup>lt;sup>8</sup> I cannot resist mentioning that I do not think that the hard problem could arise for someone who has no conception of another being that could have conscious states. But I won't pursue that here.

<sup>&</sup>lt;sup>9</sup> We have yet, of course, failed to identify the neural bases of any phenomenal qualities, though in some cases we appear to be getting close.

Talk of consciousness as what makes the mind-body problem "hard" can be found at least as far back as the late nineteenth century.

the tension in question. But could a solution to the hard problem be given without any appeal to a solution to the problem of the nature of consciousness—to the hard part of the mind-body problem, the hard nut of the problem? Suppose rational grounds were provided for the claim that it is a law of nature that any being in a certain neuro-scientific state, N, is in a state of phenomenal consciousness, C, and that there is no other state of phenomenal consciousness (save, of course, one's required to be in C) that being in N nomologically necessitates. If that would count as an answer to the hard problem, then an answer to the hard problem does not require an answer to the problem of the nature of phenomenal consciousness. For the claim in question is compatible with dualism, analytical functionalism, psychofunctionalism, type materialism, physicalist disjunctivism (described below), and superficialism (described below). If, however, the kind of explanation called for in the hard problem requires appeal to the nature of states of phenomenal consciousness, then, as will become apparent in due course, Block is mistaken in holding that the harder problem is "an epistemic add on" to the hard problem.

In any case, the harder problem of consciousness is supposed to be "closely tied to," but distinct from, the skeptical problem of other minds; and distinct as well from the hard problem. Moreover, it is supposed to be a problem that arises from an epistemic tension between naturalism and phenomenal realism that is distinct from the appearance of an explanatory gap. What this epistemic tension is—and thus what the harder problem is—is the main topic of the remainder of this part, part I, of the paper. In part II, I shall defend naturalist-phenomenal realism in the face of it. I shall argue there that while Block has raised a genuine issue, he has failed to identify an epistemic tension that puts rational pressure on naturalist-phenomenal realists to weaken or in any way compromise their position.

As we shall see in due course, the genuine issue that Block raises is the issue of robot consciousness, an issue that has been a topic of philosophical discussion for over fifty years. <sup>11</sup> Ultimately, our discussion will focus on robot consciousness since, as we shall see, Block holds that the possibility of certain sorts of robots poses the harder problem of consciousness for naturalist-phenomenal realism.

First, however, let us back up and ask what, exactly, the harder problem is. Unfortunately, there is no formulation of the problem that would be intelligible without considerable discussion. Block's formulation of it employs his own technical

-

<sup>11</sup> The "ayes" include Turing (1950) and MacKay (1952), (1956). The "nays" include Scriven (1953), Ziff (1959), Dennett (1978), Hill (1994), and Long (1994). Dennett's (1978) "nay," requires comment. He thinks that robots (or machines) could not, for instance, feel pain in the ordinary sense of pain. But, then again, he thinks that we do not really feel pain in that sense either. In a scientifically respectable sense of pain, however, he thinks that we feel pain and that robots could. Since I am concerned with the ordinary concept of pain, not some purely functional replacement concept of pain, Dennett counts as a "nay." Gunderson (1971) claims that whether a robot was conscious would depend on its hardware, rather than its program. Putnam (1964) claims that whether robots that are "psychological isomorphs of us" are conscious is indeterminate. Panpineau (2002) also holds that it is indeterminate whether a robot could be conscious, the indeterminacy due to vagueness in our phenomenal concepts. While in due course I shall address some issues that are discussed in these interesting papers and books, I lack the space to discuss any of these paper or books further here. (But see footnote 26 for a brief comment on Putnam 1964.)

terminology. Moreover, loaded with qualifications and distinctions, the dialectic of his discussion of the problem is enormously complex. Thus, considerable stage setting is required to present the harder problem. While providing such stage setting, I shall flag issues that will be discussed in part II, where our focus will turn to robot consciousness.

A key notion that Block appeals to in his lead-in to the harder problem is that of a *superficial functional isomorph* of us (normal human beings). A *superficial functional isomorph* of us is a being "that is functionally isomorphic to us with respect to those causal relations among mental states, inputs, and outputs that are specified by 'folk psychology'" (p.399). He says that such a being will share certain functional states with us, where a functional state is a kind of second-order state, a state such that being in it consists in being in a state with a certain causal role (p.400). The shared functional states will be the ones that are picked out by descriptions obtained via Ramsifying over the terms for states of phenomenal consciousness (e.g., 'pain') that appear in the conjunction of the platitudes of folk psychology. States with the causal roles in question realize the functional state.

A functional state is multiply realizable if and only if there are at least two distinct types of states that can play the causal role in question. Block tells us:

My strategy will be to start with the epistemic possibility of multiple realization and use it to argue for the epistemic possibility of multiple constitution. I shall then argue that the epistemic possibility of multiple constitution of phenomenal properties is problematic. (p.401)

I agree that it is epistemically possible for our superficial (i.e., folk psychological) functional organization to be multiply physically realized. I do not, however, take that to be reason to believe that states of phenomenal consciousness are multiply constituted; nor do I think that phenomenal realists, as a group, would think that it is. Phenomenal realists reject analytical functionalism. They hold that from the fact that a state, S, plays the causal role specified in the description obtained by Ramsification of a term for a mental state, M, in the platitudes of folk psychology, it does not follow that S constitutes M. They thus deny that realization entails constitution, and so deny that the fact of

<sup>13</sup> 

Block often speaks of functional properties, rather than functional states. But we can take state tokens to be property exemplifications, and thus take state types to be properties.

<sup>&</sup>lt;sup>13</sup> See, e.g., Lewis 1972 and Block 1980.

<sup>&</sup>lt;sup>14</sup> The states/properties that realize functional states/properties are typically called 'first order'. Block himself characterizes them as first-order (p.400). But he says in a footnote: "the restriction [of realizing properties] to first order properties is unnecessary" (p.400); and his cites Block 1990. In that paper, he argues that there may be no fundamental micro-level in nature, that nature may not "bottom out" at any level. That issue, however, is irrelevant to the distinction between first-order and second-order properties. The distinction between second-order and first-order properties is the distinction between role-properties and properties that fill those roles, not the distinction between macro-properties and micro-properties. In any case, nothing turns on that here.

Throughout the paper, I shall use 'physical' in essentially Block's unusually broad sense, so that a physical realization can be a micro-physical state, a chemical state, a biological state, a computational state, or a psychofunctional state (i.e., a deep functional role-state). (In part II, I recur briefly to the distinction between computational states and psychofunctional states.)

multiple realization (if it is one) entails multiple constitution. Indeed Block himself acknowledges that multiple realization fails to entail multiple constitution. He tells us: "the M-role can be multiply realized even if M is not multiply constituted" (p.401). Nonetheless, he seems to take the epistemic possibility of multiple realization to make plausible the epistemic possibility of multiple constitution. I do not; nor do I think the naturalist-phenomenal realist must. But of that, more in part II.

#### Block tells us that:

We have no reason to believe that there is any deep physical property in common to all and only the possible realizations of our superficial functional organization. (p.401)

While this claim is true if 'possible' means 'logically possible', it is controversial if the kind of possibility in question includes nomological possibility. Moreover, as we shall see in a moment, Block appeals to the alleged nomological possibility in question. For the sake of argument, however, I shall concede that we have, at present at least, *no* reason to believe that there is any deep physical property in common to all and only nomologically possible realizations of our superficial functional organization.

Appealing to the notion of multiple realization, Block introduces the notion of a *merely* superficial functional isomorph of us. A merely superficial isomorph is "an isomorph with respect to folk psychology and whatever is logically or nomologically entailed by folk psychological isomorphism, but that is all" (p.410). Thus, a merely superficial isomorph of us is like us in only those respects that are either logically or nomologically required for it to be a superficial functional isomorph of us. The states that realize our superficial functional organization will thus be different from the states that realize that same superficial functional organization in it; indeed Block stipulates that the realizations will be "fundamentally different" (p.405).

Moreover, while such an isomorph shares with us the superficial functional states in question, Block stipulates that it will fail to have the non-folk psychological states that a scientific psychology would attribute to us (p.410). Thus, he tells us that a merely superficial functional isomorph would not, for example, be governed by the Weber-Fechner law that just-noticeable differences in stimuli increase with increasing intensity of the stimuli, since this law is not part of folk psychology (p.410). Of course, in saying this, Block makes the controversial empirical assumption that being a deep (i.e., scientific psychological) functional isomorph of us is not nomologically required for being a superficial functional isomorph of us. But, as I indicated above, I shall spot him that assumption.

According to Block, then, a *merely* superficial functional isomorph of us shares a superficial (i.e., folk psychological) functional organization with us, but not a deep (scientific psychological) functional organization; and the superficial functional organization is realized in it in a "fundamentally different" way from the way that it is

realized in us. Suffice it to note that I shall not challenge the nomological possibility of merely superficial functional isomorphs.

There is, Block notes, a reason to think that a merely superficial functional isomorph is phenomenally conscious: namely, it is a superficial functional isomorph of a normal human being. But, he claims, given phenomenal realism, this reason is inconclusive: it can be defeated. As he puts it:

Superficial functional equivalence to us is a defeasible reason for attributing consciousness; that is, superficial functional equivalence to us provides a reason for thinking a being is conscious, but that reason can be disarmed or unmasked, its evidential value cancelled. (p.402)

Indeed he presents two considerations that could defeat functional isomorphism as a reason for attributing consciousness. Let us consider them in turn.

The first consideration that Block claims could have this defeating effect is that the merely superficial isomorph is "a homunculi head": that is, that it is such that the basis of its superficial functional organization consists of radio communications among conscious beings that stand in a one-to-one relationship with our neurons, their communicative interactions preserving the relevant inhibitory and excitatory relationships among neurons. According to Block, were a superficial functional isomorph is a homunculi head, that would defeat its functional isomorphism as a reason for attributing consciousness to it.

Let us turn, then, to the second consideration. Block says:

Here is a case in which the epistemic value of functional isomorphism is cancelled: the case involves a partial physical overlap between the functional isomorph and humans. Suppose that there are real neurophysiological differences of kind—not just complexity—between our conscious processes and our unconscious—that is, nonphenomenal—processes. Nonphenomenal neural processes include, for example, those which regulate body temperature, blood pressure, heart rate, and sugar in the blood—brain processes that can operate in people in irreversible vegetative coma. Suppose...that we find out that *all* of the merely superficial isomorph's brain states are ones that—in us—are the neural bases only of *phenomenally unconscious states*. For example, the neural basis of the functional analog of pain in the merely superficial isomorph is the neural state that regulates the pituitary gland in us. This would not *prove* that the isomorph is not phenomenally conscious (since the contexts of the neural realizers are different), but it would cancel or at least weaken the force of the reason for

8

<sup>&</sup>lt;sup>16</sup> Homunculi heads were introduced in Block (1980). It may of course very well be that we are warranted in holding that homunculi heads are nomologically impossible, and are thus ruled out by Block's condition that superficial functional isomorphs must share with us whatever is *nomologically* required for being a superficial functional isomorph of us; but, anyway, nothing in what follows will turn on this.

attributing consciousness provided by its functional isomorphism to us (pp.402-03).

Call this sort of case 'a partial physical overlap case', and the superficial functional isomorphs that fall under such cases, 'partial physical overlappers'.

I agree with Block that either of these considerations would defeat superficial functional isomorphism as a reason for attributing consciousness. Indeed that is a point I shall return to in part II. For now, however, it is worthwhile noting that these sorts of cases of merely superficial functional isomorphs— homunculi head cases and partial physical overlap cases—are standard counterexamples to the analytical functionalist's claim that superficial functional isomorphism analytically suffices for the possession of states of phenomenal consciousness. These cases show that analytical functionalism fails to capture our considered judgments about what beings are, and what beings are not phenomenally consciousness. Given that this brand of functionalism purports to be analytic, that consideration is decisive against it.<sup>17</sup>

To rule out cases of homunculi heads and partial physical overlappers, Block asks us to consider a kind of merely superficial functional isomorph that meets the following conditions: (1) the physical realizers of its "functional analogs of conscious states do not overlap with any of our brain mechanisms in any properties that we do not also share with inorganic entities that are uncontroversially mindless, for example, toasters" (p.403); and (2) it "does not have any part that itself is a functional isomorph of us and whose activities are crucial to maintaining the functional organization of the whole" (p.404). Condition (1) is intended to exclude partial physical overlappers, <sup>18</sup> and condition (2) is intended to exclude homunculi-heads. <sup>19</sup>

<sup>&</sup>lt;sup>17</sup> There are of course other standard objections to analytical functionalism. See the discussion of "absent role" cases in part II.

<sup>&</sup>lt;sup>18</sup> This condition, however, is stronger than Block intends, given his pleonastic use of 'property'. (I say that his use of 'property' is pleonastic since he speaks, for instance, of disjunctive properties (p.000).) The reason is that no superficial functional isomorph will satisfy this condition. To see this, notice that no toaster would be a merely superficial functional isomorph of us. But, then, a merely superficial functional isomorphs functional analogs of our conscious states will overlap with our brain mechanisms in a property that we do not also share with toasters, namely the property of not being a toaster. It follows that no superficial functional isomorph of us (including other human beings) will meet condition (1). For present purposes, we can reformulate the condition thus: no realization of the merely superficial functional isomorphs analogs of our conscious states are physical states (i.e., microphysical, chemical, neurobiological, computational, or deep functional role states) that we are sometimes in, and that we know have no bearing on our states of phenomenal consciousness.

<sup>&</sup>lt;sup>19</sup> In a footnote, Block mentions that condition (2) "needs further refinement" (p.404) Indeed it does since it is too strong: virtually any superficial functional isomorph that is made of atoms will satisfy it. To see this, consider any superficial functional isomorph, SF, of a normal human being. Consider next some atom at its periphery. Call the being consisting of all of SF minus that atom, SF-minus. In virtually every case, SF-minus, a proper part of SF, will be a superficial function isomorph of a normal human being. Notice that even normal human beings contain (proper) parts that are functional isomorphs of normal human beings, and whose activities are crucial to the maintaining of the functional organization of the whole. Just consider the part of a normal human being that includes all of the parts of that individual minus an atom at the periphery of the individual's brain. Call the individual in question 'Unger', the atom in question 'Sparky', and the part of Unger consisting of all Unger's parts minus Sparky, 'Unger-minus'. Unger-minus

Acknowledging that there are no doubt other sorts of realizations that would defeat superficial functional isomorphism as a reason for attributing phenomenal consciousness, Block goes on to say:

It would help if I could think of all the realizations that have these kinds of significance. If you tell me about one I have not thought of, I shall add a condition to rule it out. (p.404)

It would help because his intention is to place conditions on a merely superficial functional isomorph's realizations of its functional analogs of conscious states that ensure that those realizations are not such as to defeat superficial functional isomorphism as a reason for attributing consciousness. He is thus concerned with merely superficial functional isomorphs whose realizations fail to have that kind of significance. What kind of significance that is, exactly, is a question that I shall return to in part II.

For vividness, Block asks us to consider Commander Data, an android of *Star Treck* fame. He stipulates that Data is "a merely superficial isomorph of us (ignoring his superior reasoning and inferior emotions)" (p.402). To ensure that Data is not a partial overlapper, we are to assume that "the physical basis of Commander Data's brain is to be found in etched silicon chips rather than the organic carbon basis of our brains." (p.404). And we are to assume that Data is not a homunculi head.<sup>20</sup> Indeed Block stipulates that, "Commander Data has a realization that cannot be seen to defeat the attribution of consciousness to him either a priori or on the basis of a theory of human consciousness" (p.404). Suffice it to note that I shall discuss this stipulation in detail in part II.

Block is not as clear as he might have been about whether his stipulation that Commander Data is a merely superficial functional isomorph of us is supposed to entail that Commander Data has mental states such as beliefs and desires, or whether we are to assume that the occurrence of such terms as 'belief' and 'desire' in the platitudes of folk psychology are also to be Ramsified over in the characterization of superficial functional states. But that won't matter here. Suffice it to note that, as a phenomenal realist, I agree with Block that having beliefs and desires, even having higher-order beliefs and desires, fails to suffice (a priori) for having phenomenal consciousness. So even if we are

is a superficial functional isomorph of Unger. So Unger fails to satisfy (2). But we would of course deem Unger, a normal human being, conscious. How, then, can (2) be revised to do the intended work? For present purposes, we can simply rely on our intuitions about what sorts of beings count as Blockean homunculi-heads, and reformulate (2) as follows: it is not a homunculi-head. (Unger is not a homunculi-head despite containing Unger-minus since, for one thing, Unger-minus is not phenomenally consciousness, even though it has all that it takes intrinsically to be phenomenally conscious. 'Being phenomenally conscious' is, to use Sider's apt term, a border-sensitive predicate (Sider 2003). Border-sensitivity, of course, has nothing especially to do with consciousness. The predicate `is a mountain' is

<sup>20</sup> Block says that, "Commander Data does not have any part which itself is a functional isomorph of us and whose activities are crucial to maintaining the functional organization of the whole" (p.404). That, however, is false. To see this, just consider Data-minus.

border-sensitive. See McGee and McLaughlin 2000.)

10

j

supposed to assume that Commander Data has such higher-order beliefs and desires, the question of Commander Data's phenomenal consciousness would remain.

Block readily acknowledges that the fact that Commander Data has "a fundamentally different" realization of our superficial organization than we do is a reason to deny that Commander Data has phenomenal states. However, he says: "the fact that Commander Data's control mechanisms are fundamentally different is not a ground of rational belief that he has no phenomenal states." (p.405) (Recall that grounds for rational belief is "an epistemic level that is stronger than 'reasons for believing'" (p.407)). He holds, then, that we have a reason to believe that Commander Data is conscious and a reason to believe that Commander Data is not conscious. Both reasons are defeasible, and neither reason provides rational grounds for belief. Moreover, Block holds, we do not now know whether a hypothetical robot like Commander Data would or would not be conscious; indeed Block holds a much stronger claim that I will state shortly.

Of course, if we had a theory of the nature of phenomenal consciousness—and so, a theory that purported to crack the "hard-nut" of the mind-body problem—we could apply it to the case of Commander Data to determine whether or not Commander Data is conscious. <sup>21</sup> But while Block is a naturalist, he holds that we lack such a theory. I agree that we do not now have such a theory. And I shall not challenge the claim that we do not now *know* whether a hypothetical robot like Commander Data would or would not be phenomenally conscious.

After these preliminaries, Block offers the following "first-pass" formulation of the harder problem:

Naturalism says the default is that Commander Data is not conscious, but phenomenal realism says that the issue is open in the sense of no rational ground for belief either way. This is a first pass at saying what the harder problem. (p.407)

By saying that naturalism says the default is that Commander Data is not conscious, Block means that naturalism says that as concerns whether Commander Data is conscious, the only view for which background beliefs provide rational grounds for tentative belief is that Data is not conscious. (Recall that: "A view on a given subject is the default if it is the only one for which background considerations give rational ground for tentative belief." (p.393)) But if that is what naturalism says, then naturalism entails that we have grounds for rational belief that Commander Data is not conscious; for a tentative belief, you will recall, is a belief non-dogmatically held, and thus a belief. If, then, phenomenal realism says that the issue is open in the sense of there being no rational ground for belief either way, naturalism and phenomenal realism are not just in tension, they are incompatible. The naturalist-phenomenal realist would be committed to the view that we have rational grounds for belief that Data is not conscious and that it is

<sup>&</sup>lt;sup>21</sup> It is epistemically possible that such a theory would leave the case undecided; but then (given that the theory is correct), the answer to the question whether Commander Data is conscious would be that it is indeterminate. I shall not pursue this line, however.

not the case that we have rational grounds for belief that Data is not conscious. Needless to say, that is not a consequence a naturalist-phenomenal realist could "learn to live with"

Neither naturalism nor phenomenal realism, however, says what Block says they say. Consider first naturalism, and then phenomenal realism.

Naturalism is default physicalism. As Block himself points out, both disjunctivism and superficialism are kinds of physicalism (in the sense in question), and are compatible with Data's being conscious (p.413). He says: "Disjunctivism is a form of physicalism that allows that consciousness is a physical state that is disjunctive by the standards of physics" (p.408). Moreover, he points out:

Disjunctivism is one way of making naturalism compatible with Commander Data's being conscious, but there are others. One is the view that consciousness is a matter of empirical fact identical to the superficial functional organization that we share with Commander Data. We might call this view superficialism...Superficialism is the phenomenal-realist claim that consciousness is identical to the superficial functional organization that we share with Commander Data—as distinct from the deflationist version of this claim mentioned earlier [i.e., analytical functionalism]. Note that superficialism says consciousness is a role property, not a property that fills or realizes that role. A role property is a kind of dispositional property. (p.412)

Thus, even given Block's stipulations about Data, naturalism fails to entail that the default is that Data is not conscious. Further assumptions must be added to naturalism to get that result.

Nor does phenomenal realism say that whether Commander Data is conscious is open in the sense of "no rational ground for belief either way." It is true that:

(i) Phenomenal realism neither entails that there is rational ground for belief that Commander Data is conscious nor entails that there is rational ground for belief that Commander Data is not conscious.

#### But it is not true that:

(ii) Phenomenal realism entails that there is no rational ground for belief that Commander Data is conscious and that there is no rational ground for belief that Commander Data is not conscious.

Neither entailing A nor entailing B does not of course entail entailing (not-A and not-B). It is true that phenomenal realism leaves open whether Commander Data is conscious. But it does so only in this sense: it neither entails that Data is conscious nor entails that Data is not conscious.

Moreover, given that phenomenal realism leaves open whether Data is conscious only in the sense just described, even if naturalism together with unimpeachable assumptions entailed that it is the default that Data is not conscious, it would not follow that naturalism and phenomenal realism are in epistemic tension. There is no epistemic tension in holding a view that leaves open whether Data is conscious (in the sense of not entailing either way), while also holding a view that entails that the default is that Data is not conscious. There is no epistemic tension in holding a view that leaves open whether P (in the sense of not entailing either way), while also holding a view that entails that one has rational grounds for the belief that not-P.

The passage just under discussion, you will recall, is merely a first pass at saying what the harder problem is. Block goes on to provide a more careful formulation of the problem to which I shall now turn.

Block divides the harder problem into a two-part problem. Here is the first part, along with a central part of his reasoning for it:

Our knowledge of physicalism is doubly problematic: [1] we have no conception of a ground of rational belief that Commander Data is or is not conscious, and [2] we have no way of moving from a conclusion that Commander Data is conscious to any consequence for the truth of physicalism. And this holds despite the fact that [3] physicalism is our default view. [4] *Physicalism is the default and also inaccessible and meta-inaccessible*. This is part of the harder problem. A second part follows. (p.413) [Emphasis Block's; the inserted numerals are mine.]

Block's italicized claim, which I have labeled '[4]', is the first part of the harder problem. Claims [1], [2], and [3] are part of the reasoning that leads to [4]. I'll discuss these claims in reverse order.

By saying that [4] is the first part of the harder problem, there is some uncertainty as to whether Block means that the first part of the harder problem for naturalist-phenomenal realists is that [4] is true, or whether he means that the first part of the harder problem for them is that they are committed to [4]. This uncertainty is due to the fact that Block argues for [4], and seems to take it to express a genuine epistemic tension, albeit one that a naturalist-phenomenal realists could nevertheless "learn to live with." But when Block argues for [4], he does so from the perspective of a naturalist-phenomenal realist. (He is of course himself one.) And it is clear that he thinks that naturalist-phenomenal realists are committed to [4]. For reasons that I shall state shortly then, it is best to take the first part of the harder problem for naturalist-phenomenal realists to be that they are committed to [4], rather than that [4] is true.

Claim [4] claim is, of course, a conjunction, the first conjunct of which—physicalism is the default—means that as concerns phenomenal consciousness, physicalism is the only view for which background considerations provide rational grounds for tentative belief. The second conjunct employs Block's technical notions of

inaccessibility and meta-inaccessibility. Here is how he introduces the notion of meta-inaccessibility:

We have no conception of a ground of rational belief to the effect that a realization of our superficial functional organization that is physically fundamentally different along the lines I have specified for Commander Data is or is not conscious. To use a term suggested by Martine Nida-Rumelin in commenting on this paper, Commander Data's consciousness is meta-inaccessible. Not only do we lack a ground of belief, but we lack a conception of any ground of belief. This meta-inaccessibility is a premise... (p.405)

While Block fails to explain the notion of inaccessibility, his text gives the impression that the notion is that of having no rational grounds for the belief or the disbelief in question. If that is what it means, then [4] means: as concerns phenomenal consciousness, physicalism is the only view for which background considerations provide rational grounds for tentative belief, and also we have no rational grounds for belief or disbelief in physicalism and no conception of how we could have rational grounds for belief or disbelief in physicalism. But notice, then, that [4] entails a contradiction: we have rational grounds for belief in physicalism and it is not the case that we have rational grounds for belief in physicalism. For if background considerations provide us with rational grounds for tentative belief in physicalism, then we have rational grounds for belief in physicalism. But [4] also entails that we have no rational grounds for belief in physicalism. Thus, it entails a contradiction.

Since Block does not purport to claim to have shown that naturalism and phenomenal realism are incompatible, perhaps 'inaccessible' should not be taken to mean 'no grounds for rational belief or disbelief." What 'inaccessible' means, then, is uncertain, and so what [4] means is uncertain. In any case, as we saw, 'physicalism is metainaccessible' means by stipulation that we have no conception of how we could have rational grounds for belief in physicalism. Given that, [4] entails that as concerns phenomenal consciousness, physicalism is the only view for which we have rational grounds for belief and we have no conception of how we could have rational grounds for belief in physicalism. This claim could be true. But notice that this claim is one we could never be justified in believing or asserting.

On either reading of [4], then, commitment to [4] is not something with which one can learn to live. For on one reading [4] entails a contradiction; on another it is a claim that one could never be justified in believing or asserting. By saying that [4] is the first part of the harder problem, then, we shall take Block to mean that naturalist-phenomenal realists are committed to [4] (whatever exactly [4] means). Thus, it won't due for a naturalist-phenomenal realist to respond to Block by simply rejecting [4]. Naturalist-phenomenal realists must respond to his argument that they are committed to [4].

Claims [1], [2], and [3] are part of the reasoning that purportedly leads to [4]. A further premise in the reasoning, unmentioned in the passage under discussion, is: [3.5]

<sup>&</sup>lt;sup>22</sup> An epistemic internalist, though, might dispute that.

that we have no conception of how we could have rational grounds for believing that either disjunctivism or superficialism is true (p.415) Now [1], [2], [3], and [3.5] fail to jointly entail [4]. However, Block seems to take these claims to make plausible the claim that naturalist-phenomenal realists are committed to [4].

As will become apparent in part II, I am in sympathy with Block's claim [3.5].<sup>23</sup> Moreover, naturalist-phenomenal realists are, of course, committed to [3] by their naturalism. Further, [2] is true since disjunctivism and superficialism are types of physicalism, and they do not entail that Commander Data is not conscious. Let us, then, turn to the remaining premise, [1], namely, that we have no conception of a ground of rational belief that Commander Data is or is not conscious.

In claiming [1], it might appear that Block has in mind an instance of the skeptical problem of other minds. In its most general form, the problem of other minds is the problem of how one could know (/have rational grounds for believing/be justified in believing, etc.) that anything other than oneself now has or lacks a mind. It may thus appear that the question of how one could have rational grounds for believing or for disbelieving that Commander Data is conscious is just a special case of this problem. But Block says, you will recall, that while the harder problem "is tied closely to the problem of other minds," (p.392) it is not a skeptical problem. He says: "I will be adopting a point of view that sets skepticism aside. *Undoubtedly*, humans are conscious and rocks and laptops are not." (p.399) (Emphasis his.) Moreover, he says, you will also recall, that grounds for rational belief is "an epistemic level that is stronger than 'reasons for believing' and weaker than 'rational certainty'. I take it that a ground of rational belief that p allows knowledge that p but mere reason for believing that p does not" (p.407). In requiring something weaker than rational certainty for grounds for rational belief, Block means to be eschewing the standards that a philosophical skeptic might maintain are required for knowledge.

As I indicated at the outset of the paper, however, Block says nothing at all about what it takes to know something. Indeed despite the fact that he invokes the notion of knowledge in his characterization of grounds for rational belief, he actually seems not to be especially concerned with knowledge. The real issue that engages him is whether we could be in a position to reasonably believe or be warranted or justified in believing that Commander Data is or is not conscious, not by the skeptic's standards of 'reasonable belief' or 'gustified belief', but by the standards actually employed in the practice of science; more specifically, by those standards actually employed in the practice of special sciences such as, for example, the life sciences. How, by such scientific standards of reasonableness, could one be in a position reasonably to conclude that Commander Data is or is not conscious? This is, I believe, the question that Block means to pose. And this is the question that concerns me. For I am here, like Block,

\_

<sup>&</sup>lt;sup>23</sup> Block also makes a stronger claim about disjunctivism and superficialism, namely that they are both meta-inaccessible (p.000); that is, that we have no conception of how we could have rational grounds for belief that they are true and no conception of how we could have rational grounds for belief that they are false. I disagree with this stronger claim. As will be clear in part II, I think we have a conception of how we could have rational grounds for believing disjunctivism and superficialism are false.

concerned with the prospects for a reductive scientific theory of phenomenal consciousness, rather than the prospects for an answer to the skeptical problem of other minds.

Block explicitly notes that while he maintains [1], he does not maintain, as does Colin McGinn (1989), that we are *cognitively closed* to the physical conditions that make it the case that a being is phenomenally conscious (p.406). He says: "I am not denying that we might someday come to have the conception we now do not have. (So I am not claiming—as McGinn does—that this knowledge can be known now to be beyond our ken.) I am merely saying that, at this point, we have no idea of evidence that would ground rational belief, even a hypothetical or speculative conception" (p.406) And he says: "I do not deny that one day the question of whether a creature like Commander Data is phenomenally conscious may *become* a testable empirical claim. But it is obvious that we do not *now* have any conception of how it could be tested." (p.406) [Emphases Block's.]<sup>24</sup>

Why, then, does Block think that we *now* have no conception of how we could have grounds for rational belief that Data is or is not consciousness? As concerns [1], he says:

This...is a premise rather than a lemma or a conclusion because the line of thought I have been presenting leads up to it without anything that I am happy to think of as an argument for it. My hope is that this way of leading up to it will allow the reader to see it as obvious. (p.405)

His way of leading up to claim [1] was recapitulated above.

Despite Block's hope that readers will find [1] obvious, I think that [1] is false. We have a conception, a speculative, hypothetical conception, of how we could have grounds for rational belief that Commander Data is not conscious. Moreover, we have a conception of how we could have such grounds without naturalism or phenomenal realism being at all compromised. As Block notes, while "phenomenal realism rejects...armchair, philosophical-reductive analyses" (p.392) of consciousness, "phenomenal realists have no brief against scientific reduction of consciousness." (p.392). Naturalist-phenomenal realists, I contend, can argue that we have a conception of how we could have rational grounds for a scientific reduction of consciousness that entails that Commander Data not conscious. Naturalist-phenomenal realists can thus respond to the first part of Block's harder problem by denying, on the grounds that [1] is false, that they are committed to [4]. Or so I'll argue in part II of this paper.

Let us turn, then, to Block's full statement of the harder problem:

The first of the epistemic difficulties...is that physicalism is the default, but also inaccessible and meta-inaccessible. We are now ready to state the second

16

Needless to say, Block no doubt means to exclude as irrelevant such tests as that an Erwin-type "confirmation machine" might print out that Commander Data is not conscious (Erwin 1970).

epistemic difficulty. Let us introduce a notion of the *subjective default* view which we have rational ground for believing on the basis of background information—but only ignoring escape hatches—such as disjunctivism and superficialism—which we have no evidence against but which are themselves inaccessible and meta-inaccessible. Then the second epistemic difficulty is that of holding *both that it is an open question whether Commander Data is conscious and that it is the subjective default view that he is not.* These two epistemic difficulties constitute the harder problem. (p.418) (Emphasis his.)

One could not rationally hold both that it is an open question whether Commander Data is conscious and that one has rational grounds for believing that Commander Data is not conscious. One could, however, rationally hold that that it is an open question whether Commander Data is conscious and that it is the subjective default that Commander Data is not conscious. If one holds that it is the subjective default that Commander Data is not consciousness, then one holds that *given background beliefs*, *but ignoring escape hatches*, *one has rational ground for believing that Commander Data is not conscious*. Holding that it is the subjective default is thus compatible with holding that one lacks rational grounds for believing that Commander Data is not conscious. For one might hold that the escape hatches cannot properly be ignored and cannot be ruled out, and so not believe that one has rational ground for believing that Commander Data is not conscious.

Of course, if naturalist-phenomenal realists are committed to holding that they have rational grounds for believing that Commander Data is not consciousness, and committed as well to holding that they cannot properly ignore or rule out hypotheses that entail that Commander Data is conscious, then naturalist-phenomenal realism is indeed in epistemic tension: it is incoherent. But while phenomenal realism itself leaves open whether Commander Data is conscious, we have a conception of how a naturalist-phenomenal realist could be in a rational position to deny that Data is conscious without compromising his or her naturalism or phenomenal realism. Or so, as I have already indicated, I shall argue in arguing that [1] is false.

In a nutshell, then, my naturalist-phenomenal realist response to Block's two-part harder problem will be that we have a conception of how we could have rational grounds for believing that Commander Data is not conscious that involves no compromise of naturalism or phenomenal realism. In arguing this, I shall not argue that the combination of naturalism and phenomenal realism is free of any epistemic tension. I think, for instance, that the appearance of an explanatory gap reveals an epistemic tension. But the harder problem is supposed to be different from the familiar problems that reductive physicalist solutions to the hard problem face; it is not supposed to collapse into some familiar problem for physicalist solutions to that problem (of which more in part II). And so, I shall take it that the issue of Commander Data's consciousness is supposed to pose a problem for naturalist-phenomenal realism that is different from, and additional to, the familiar problems in question. It is this different problem that will be my primary concern.

As promised earlier, the issue that confronts us boils down to the one of robot (or alien being) consciousness. I want to underscore that it boils down to the issue of robot phenomenal consciousness. What is at issue is not whether we have a conception of how we could have rational grounds for denying that Commander Data actually speaks English, or for denying that that Data has higher-order propositional attitudes, or for denying that Data has propositional attitudes. Nor is the issue whether there is a sense in which Data can, for instance, see or hear. There is a straightforward sense in which gold fish can see, but there is a serious question indeed of whether they have visual experiences, or indeed any phenomenal experiences at all. (There may be something that it is like to be a bat, but is there anything that it is like to be a gold fish?) The issue is whether we have a conception of how we could have rational grounds for denying that Data is phenomenally conscious, that is, for denying that it is like something for Data to be in various states. Someone might wish to argue that if Data fails to be phenomenally conscious, then Data fails to actually speak English, to have higher-order propositional attitudes, or to have propositional attitudes, or that there is no sense in which Data can see or hear, etc. I, however, would not argue for any of those claims since I don't believe any of them.<sup>25</sup> The issue is just whether we have a conception of rational grounds for denying that Data is phenomenally conscious—for denying that Data has what Block calls phenomenality (p.394).<sup>26</sup>

Before pursuing this issue, however, it is worthwhile noting that one might sympathize with Block's view that we have no conception of how we could have rational grounds for believing or disbelieving that Commander Data is conscious. And one might think that itself reveals an epistemic tension between naturalism and phenomenal realism. Perhaps then, one might think, the harder problem for the naturalist-phenomenal realist is really just that we have no conception of how we could have rational grounds for believing that Commander Data is or is not conscious.

But note that even if we lack any such conception, that would not be due to our embracing naturalist-phenomenal realism. For compromising or weakening naturalism

.

<sup>&</sup>lt;sup>25</sup> I do not deny, however, that possession of phenomenal consciousness is required for full possession of certain concepts, and for full mastery of certain words of English. I think that this is the case, for instance, for the concept of pain and the word 'pain'. (See McLaughlin 2003.) Full mastery of that term requires the ability to apply it one's states of pain by introspection. But it is, I believe, possible for a being that is devoid of phenomenal consciousness to engage in fluent conversation using 'pain', and to have a partial grasp of the concept of pain, if it is a member of a linguistic community that has masters of the concept of pain. (Hilary Putnam (1964) was thus mistaken when he claimed that the issue concerning robot consciousness was whether we should decide to count them as members of our linguistic community. Even if we so counted them (and they so count themselves), the issue of their consciousness would remain.) I think that it is possible for such a being to believe—mistakenly of course—that it is in pain. One can exercise in thought concepts of which one lacks a complete grasp.

Would I allow that Commander Data can *think* even if Commander Data cannot *feel*? I would not if by 'thinking' one means having an occurrent thought with a linguistic image of some sort. For such occurrent thoughts are or at least involve states of phenomenal consciousness. But I shall at least leave open whether Commander Data can think in the sense of reason, even if Commander Data cannot feel. (As I mentioned earlier, Block may have intended his stipulation that Commander Data is a superficial functional isomorph of us to entail that Commander Data can reason. It is uncertain since we were not told whether nonphenomenal terms of folk psychology were to be Ramsified. But as I also mentioned, that won't matter for present purposes since we are assuming phenomenal realism.)

would not help at all. Suppose that we flat-out rejected naturalism, thereby denying that consciousness has a scientific nature. Suppose that we embraced dualism, maintaining that types of states of phenomenal consciousness are ontologically fundamental, irreducible states. That would do nothing to enable us to decide the issue of Commander Data's consciousness.

If, however, we have no conception of how we could have rational grounds for believing that Commander Data is or is not conscious, would not that be due to our embrace of phenomenal realism? No, it would have nothing whatsoever to do with our embrace of phenomenal realism. There *is no* a priori sufficient condition for phenomenal consciousness that can be stated (noncircularly) in nonphenomenal terms. *Replacing* our concepts of consciousness with "deflationary" purely functional concepts that have such a priori sufficient conditions (something that I do not believe our cognitive architecture would permit in any case) would not change that. It would only change the subject. Moreover, to deny that our concepts of consciousness (our phenomenal concepts) are actually applicable to us—to claim that only deflationary counterparts of them are actually applicable to us—would be to deny that we are conscious. That is of course one possible response to the case of Commander Data: Commander Data is not conscious, but neither are we. Suffice it to note that I myself cannot see why anyone who had the sense that the issue of Commander Data's consciousness is rationally undecidable would respond by rejecting phenomenal realism.

But I cannot see a lot of things. One thing that I think I can see, however, if only through a dark glass dimly is

# II. How A Naturalist-Phenomenal Realism Could Rationally Decide That Commander Data Is Not Conscious.

In a later, expanded version of "The Harder Problem of Consciousness" that Block has posted on the web, he makes the following helpful remarks:

We could see the [harder] problem this way: if Commander Data is conscious, then we have a choice of superficialism, disjunctivism and dualism. The naturalist wants to reject dualism, but it is cold comfort to be told that the only alternatives are doctrines that are epistemically inaccessible. So this may lead us to want to say that Commander Data is not conscious. But we have no evidence that he is or is not conscious.

(www.nyu.edu/gsas/dept/philo/faculty/block/papers/harder.htm, p.31)

There are theories of mind other than superficialism, disjunctivism, and dualism that are compatible with Commander Data being conscious: for example, analytical functionalism, panpsychism, and neutral monism, to name a few. But I take Block's essential point to be that if we embrace naturalist-phenomenal realism, then if Commander Data is conscious, we have a choice of superficialism or disjunctivism. As will be apparent, I agree with Block that superficialism and disjunctivism would be cold

-

<sup>&</sup>lt;sup>27</sup> I shall recur to the topic of deflationary, purely functional concepts in part II.

comfort indeed. As will also be apparent, I think that he is right that "we have no evidence that [Data] is...conscious." I think, however, that he is mistaken when he says: "we have no evidence that [Data] is...not conscious," if by 'evidence', he means an undefeated epistemic reason. The fact that we are left with only unpalatable theoretical consequences if Commander Data is conscious is itself an epistemic reason for believing that Commander Data is not conscious; and it is, I believe, one that can withstand defeat.

If Commander Data is conscious, then the options for naturalist-phenomenal realists are even narrower than Block claims. If Commander Data is conscious, then the only option left for a naturalist-phenomenal realist is disjunctivism. I say that because, philosophical skepticism aside, we already know that superficialism is false. We know that because we know that superficial functional states are not even nomologically required for states of consciousness. As Block notes in dismissing the skeptical problem of other minds: "Undoubtedly, other humans are conscious and laptops are not" (p.399) (Emphasis his.) And he adds in the later web version: "Further, bats are undoubtedly conscious" (p.8) (Emphasis his.) Babies and bats are conscious, yet they are not superficial functional isomorphs of normal adult human beings. We would, moreover, regard ourselves as justified in believing that a fellow human being suffering from paralysis and severe Alzheimer's disease was in pain, if we knew that the individual was in a physical state that pain accompanies in normal individuals. We would so regard the individual, even though the individual's state would not play the causal role that folk psychology associated with being in a pain state. These familiar absent role cases make a powerful case that states of consciousness fail to even nomologically require superficial functional states, and so that superficialism is false. We see, then, that if naturalistphenomenal realism is true and Data is conscious, then disjunctivism is true.

Disjunctivism is riddled with problems. As Block mentions, it runs counter to our intuition that consciousness is a unified phenomenon (p.000). Moreover, there are problems of causal efficacy: how could the disjunctive properties endow anything with causal powers? Block argues at length that disjunctivism is "epistemically inaccessible" on the grounds that the a posteriori claim that a state of conscious is identical with a heterogeneous disjunction of physical states could not explain the nomological correlation between the disjunctive state and the state of consciousness. I demur on that point. The identity claims, if true, would be metaphysically necessary, and so would entail the nomological correlations. Moreover, if such a identity claim were true, then it seems that the claim would explain why being in one of the disjunct states (metaphysically) suffices for being in the conscious state in question: it does so since it suffices for being in the disjunctive state, and that state *is* the conscious state. I am in great sympathy, however, with Block's claim that we have no conception of how we

<sup>&</sup>lt;sup>28</sup> Block does not say what he means by a *heterogeneous* disjunctive property. He may mean a disjunction of properties that are not determinates of some determinable property. In any case, I shall use 'heterogeneous disjunctive property' in that way. But I won't venture an account of the determinable distinction.

As Kripke (1980) showed, given the indiscernbility of identicals and the fact that everything is such that it is necessarily identical with itself, we can derive: (x)(y) (if x = y, then necessarily x = y).

could be justified in inferring such identity claims as the best explanation of such correlations. I shall return to that issue shortly.

Why, then, think that Commander Data is conscious? There is, you will recall, one, and only one, reason: Data is a superficial functional isomorph of a normal human being. The reason is defeasible. As we saw, the fact that a superficial functional isomorph is a homunculi head or a partial physical overlapper would defeat this reason for attributing consciousness. Of course, Block stipulates that Commander Data is not a homunculi head or a partial physical overlapper. Indeed he stipulates, you will recall, that:

Commander Data has a realization that cannot be seen to defeat the attribution of consciousness to him either a priori or on the basis of a theory of human consciousness. (p.404)

Being a homunculi-head and being a partial physical overlapper, however, do not defeat superficial isomorphism as a reason for consciousness a priori: they defeat it based on what we know about human consciousness. To see this as concerns cases of partial physical overlappers, recall that our reason for thinking that they are not conscious is that we know the physical states that realize their superficial functional organization are ones that in us have no bearing on consciousness. And as concerns homunculi-heads, we of course know that we are not homunculi heads. But were we to come to believe that we are homunculi-heads would not we change our verdict on the issue of homunculi head consciousness? To consider a distant but less remote doxastic situation, suppose we discovered that we are androids with a silicon-based material composition and structure of the sort we imagine Data to have. We would, then, I claim, take ourselves to have rational grounds for believing that Data is conscious.

Block can stipulate that Commander Data has a realization that cannot be seen to defeat the attribution of consciousness a priori. He cannot, however, stipulate that Data has a realization that cannot be seen to defeat the attribution of consciousness on the basis of a theory of human consciousness without begging the very question at issue.

Consider why it is that we think: "*Undoubtedly*, other humans are conscious and laptops are not" (p.399) (Emphasis his.) As Block notes, "phenomenal realism is based on one's first-person grasp of consciousness" (p.392) Indeed we know from our first-person perspective that we are conscious. Moreover, we assume that we are justified in attributing consciousness to other beings if and only if they are relevantly like us. We take our fellow human beings, members of our own species, to be relevantly like us (and members of neighboring species to be relevantly like us as well); and we take laptops to be relevantly unlike us. So we take ourselves to be justified in attributing consciousness to our fellow human beings, and to be justified in denying that laptops are conscious. That our concepts of consciousness are in this way anthropocentric is the insight behind John Stuart Mill's argument from analogy.<sup>31</sup>

See Hill 1991, Chapter 9.

The best defense of the argument from analogy to date is in Hill 1991, Chapter 9.

Of course, everyone would agree that a being is conscious if and only if it is relevantly like us, if being relevantly like us includes having phenomenal consciousness. However, naturalist-phenomenal realists hold that a being is conscious if and only if it is relevantly like us in physical respects. <sup>32</sup> They do not regard this as a priori. They allow, for instance, that (philosophical) zombies, beings that are like us in every respect save having consciousness, are coherently conceivable. Moreover, they allow that it is coherently conceivable that laptops are conscious. However, it is compatible even with dualism that it is nomologically necessary that a being is conscious if and only if it is relevantly like us in physical respects. Indeed property dualists who eschew substance dualism standardly hold that conscious states are nomologically correlated with physical states (at least in our broad sense of 'physical'). <sup>33</sup> And naturalists-phenomenal realists, of course, hold that despite being a posteriori, it is metaphysically necessary that a being is conscious if and only if it is relevantly like us in physical respects. They hold this since they hold that conscious states are identical with physical states.

### Block observes at one point:

Elliot Sober's "Evolution and the Problem of Other Minds" argues plausibly that our rationale for attributing mental states to other humans is a type of 'commoncause' reasoning. But such common-cause reasoning is vulnerable to evidence against a common cause, for example, evidence for lack of genealogical relatedness or evidence for different scientific bases for similarity of behavior that is exhibited. Thus the rationale for attributing mentality to humans does not fully apply to Commander Data. (p.422)

I think that we already have rational grounds for thinking that it does not apply at all to Commander Data. Superficial functional isomorphism to us is evidence for phenomenal consciousness only insofar as it is evidence for sameness of realization. And by stipulation, Data's superficial functional organization has a fundamentally different realization from ours. That is why we indeed "have no evidence that [Data]...is conscious" (www.nyu.edu/gsas/dept/philo/faculty/block/papers/harder.htm, p.31).

Of course, it may be that Data is conscious, despite that. It is epistemically possible that Data's different realization necessitates consciousness, and that consciousness is a causal factor (albeit an independent overdetermining one) vis-à-vis the relevant causal effects. I am not denying that, for it does not tell against my claim that superficial functional isomorphism is evidence for consciousness only insofar as it is evidence for sameness of realization. We know that superficial functional isomorphism

<sup>&</sup>lt;sup>32</sup> Block notes at one point that he is in agreement with John Searle (1992) "that whether physically different realizations of human functional organization are conscious is not an a priori matter and could be said to depend on whether their brains have 'equivalent causal powers' to ours—in the sense of having the power to be the physical basis of conscious states" (p.406). But that is not the point. No one would dispute the claim that whether a creature is conscious depends on whether it possesses something with the causal powers of our brains, if those powers include the power to be the physical basis of conscious states.

Indeed even substance dualists look for nomological covariations (though not nomological conistantiations) between states of phenomenal consciousness and physical states. Recall Descartes and the pineal gland.

fails to suffice for consciousness: homunculi-heads and partial physical overlappers are superficial functional isomorphs, yet are not conscious. When, however, superficial functional isomorphism is due to sameness of realization, then it is evidence for consciousness. The reason is that we know that our own realization (whatever it is) is accompanied by consciousness. The fact that Data's realization of our superficial functional isomorphism is fundamentally different from ours thus defeats Data's superficial isomorphism as a reason for attributing consciousness to Data. Given the fundamental difference in realization, superficial functional isomorphism is not a reason for attributing consciousness.<sup>34</sup>

The fact that the only reason for attributing consciousness to Commander Data is defeated does not of course entail that Commander Data is not conscious. In trying to develop a scientific theory of consciousness, however, should we take care not to close off the question of Commander Data's consciousness? We could avoid closing off the question only by embracing disjunctivism. Should we, then, in trying to develop a scientific theory of consciousness, embrace disjunctivism—the view that conscious states are identical with disjunctions of physical states—, despite all of its problems? Or would that be to take the possibility of Commander Data's consciousness too seriously?

Human babies are undoubtedly conscious. They share some superficial functional similarities to adults. Suppose that we someday develop silicon-based robots that are superficial functional isomorphs of babies. Should we take seriously the possibility that their realization metaphysically suffices for consciousness? In trying to develop a scientific theory of consciousness, should we take care not to close off the question of robot-baby consciousness? Bats, Block says, are undoubtedly conscious. Suppose, then, that the Pentagon someday deploys silicon-based robot bats that are superficial functional-isomorphs of bats. Should a scientific theory take care not to close off the question of their consciousness? We regard rabbits are conscious. Someday the shelves of the toy stores may be filled with silicon-based robot rabbits that shared essentially as much with us by way of superficial functional organization as do rabbits. Should a scientific theory take care not to close off the question of their consciousness?

Consider, again, Commander Data. Suppose that Commander Data becomes paralyzed as a result of damage. Suppose further that as a result of the damage, Commander Data's capacities diminish to the point that Data bears no closer resemblance to us in superficial functional organization than does an adult, human paralytic Alzheimer's sufferer. We would, then, fail to share any physical properties with Commander Data save heterogeneous disjunctive physical ones. (And, of course, we share heterogeneous disjunctive physical properties with laptops and rocks, indeed with

<sup>&</sup>lt;sup>34</sup> I take the defeater here to be an undercutting defeater, in Pollock's sense (1974, 1986); to be a secondary defeater in Field's sense (1996). (Pollock's undercutting defeaters are Field's secondary defeaters; and Pollock's rebutting defeaters are Field's primary defeaters.) Thus, I think that we have *no* evidence that Commander Data is conscious. One could, however, maintain that it is, rather a rebutting defeater (or a primary defeater). The issue will turn on epistemological issues that won't really matter for present purposes. If one takes the defeater to be a rebutting defeater, then the point is that the little evidence that Commander Data's superficial isomorphism provides for the belief that he is conscious is considerably outweighed by his difference from us in the realization of that functional organization.

anything that is physical.) Suppose, then, that Commander Data is in the physical state, P, that once played the superficial role of pain, but that is now diminished in that role to the same extent that we might find in a human paralytic Alzheimer's sufferer. Should we take seriously the possibility that Commander Data is in pain? Suppose that an intact, proper functioning silicon-based robot with a smooth round-cylinder shape and a limited range of behavioral dispositions were such that whenever it is in P, it spins slowly in a circle. Or suppose that it is always in P, no matter which of the fairly simple activities it engages in. Should a scientific theory take care not to close off the possibility that P metaphysically suffices for being in pain? Of course, as to whether it metaphysically suffices, the philosophical skeptic will say: you never know. But it is hard to see why a scientific theory of consciousness should take this possibility seriously; or indeed any of the possibilities mentioned in the preceding paragraph.

Given that if Commander Data is conscious, naturalist-phenomenal realists would be left with only the unpalatable option of disjunctivism, they might well think that I have now already said enough to provide the material to make a case that we have a conception of how we could have what would count as rational grounds in the special sciences for denying that Commander Data is conscious. We have no (undefeated) reason to believe that Commander Data is conscious. If we regard Commander Data as conscious, we are left with disjunctivism. Reasons of theoretical simplicity and overall coherence provide rational grounds for rejecting disjunctivism. Since naturalist-phenomenal realism is true, and we have rational grounds for rejecting disjunctivism, we have rational grounds for denying that Commander Data is conscious.

But Block poses the case of Commander Data as a challenge to naturalist-phenomenal realism, and so I cannot rest my case here. Recall, though, that to reject phenomenal realism is just to change the subject. Moreover, it is hard to see how, for instance, rejecting naturalism in favor of dualism would help matters. Should the property dualist, who eschews substance dualism, take seriously the possibility that silicon robot babies, robot bats, and robot rabbits are conscious? Should such a dualist take seriously whether physical state P nomologically suffices for pain? In trying to arrive at laws of nomological correlation, should such a dualist take care not to close off these possibilities? Indeed should the substance dualist take care not to close off these possibilities when searching for laws of co-variation? We might of course retreat to panpsychism. But then we were wrong in thinking that laptops, rocks, and indeed quarks are not conscious.

In any case, I shall stop raising rhetorical questions, and shall try, instead, to bring out the apparent tension that arises from the issue of robot consciousness in the context of naturalist-phenomenal realism. Then, I shall respond to it.

After making his own an extended case that disjunctivism and superficialism fail to be epistemically viable, Block says:

The only epistemically viable naturalist or physicalist hypothesis—the only naturalist or physicalist hypothesis we have a conception of a reason for

accepting—is a deep unitary physical or otherwise scientific property in common to all and only conscious beings, a naturalistic basis that Commander Data does not share. (p.413)

As should be apparent, this seems to me enormously plausible. If it is right, then any epistemically viable naturalist or physicalist hypothesis will have as a consequence that Commander Data is not conscious.

There are essentially two brands of physicalism that entail that there is a deep unitary physical property in common to all and only conscious beings:

*Type Materialism*. For every state of phenomenal consciousness C, there is some neuro-scientific state, N, such that C = N. <sup>35</sup>

*Psychofunctionalism*. For every state of phenomenal consciousness C, there is some deep functional role state, F, such that C = F.

Following Block, I take functional role states to be second-order states: states of being in some state that plays a certain causal role. Deep functional role states are characterized via Ramsification of the theoretical terms of a scientific psychology. By neuro-scientific states, I mean the sorts of states that neuroscience, including computational neuroscience, might postulate. These include states of large pools of neurons, and of systems of networks of neurons. The state of one's c-fibers firing is a neuro-scientific state. Block mentions the state of cortico-thalamic oscillation, another neuro-scientific state. Neuroscientific states also include abstract neuro-computational states. Thus, consider opponent processes theory, the leading neuro-computational theory of color vision.<sup>36</sup> That theory postulates two pairs of opponent information processing channels: the red-channel and the green-channel, which are opponent channels, and the blue-channel and the vellowchannel, which are opponent channels. According to the theory, one has a visual experience of unique red (i.e., red that is not at all bluish or yellowish) if and only if one's red-channel is activated, one's green-channel is deactivated, and one's blue-channel and yellow-channel are in equilibrium. Notice that this computational state is not itself a functional role-state, even though the channels in question are characterized as consisting of mechanisms that can compute various functions, and the fact of opponency is functional-role fact (a dispositional fact). There are some differences in the literature as concerns what would sorts of states count as type materialist states and what sorts count as psychofunctional states. But as I said, following Block, I take the latter to be restricted to second-order states—to role states. In any case, for present purposes, what is important is not the difference between type materialism and psychofunctionalism, but rather what they have in common. They are brands of physicalism in Block's broad sense. And they

<sup>&</sup>lt;sup>35</sup> A type materialist about phenomenal consciousness need not, of course, hold that other sorts of mental states are identical with neuro-scientific states. A type materialist about phenomenal consciousness could be a psychofunctionalist about other sorts of mental states (e.g., beliefs, desires, and intentions), or indeed even an analytical functionalist about them.

<sup>&</sup>lt;sup>36</sup> Hurvich 1980.

both entail that Commander Data is not conscious. For, by stipulation, Commander Data fails to share any neuro-scientific or deep functional role property with us.

Now one might think that the fact that type materialism and psychofunctionalism entail that Commander Data is not conscious is a problem for these views. If there is a problem, however, it would not arise simply from phenomenal realism. As we noted earlier, phenomenal realists "have no brief against scientific reduction of consciousness" (p.392). Of course, type materialism and psychofunctionalism would face a serious problem if in fact we had no conception of how we could have rational grounds for believing that Commander Data is not conscious. For if that were so, then we would have no conception of how we could have rational grounds for believing either type materialism or psychofunctionalism. But, on the other hand, if we have a conception of how we could have rational grounds for believing either type materialism or psychofunctionalism, then we would have a conception of how we could have rational grounds for believing that Commander Data is not conscious. I shall argue shortly that we have such a conception. But first some further preliminaries are in order.

Type materialism and psychofunctionalism, of course, purport to offer solutions to the hard problem. Recall:

The hard problem is one of explaining why the neural basis of a phenomenal quality is the neural basis of that phenomenal quality rather than another phenomenal quality or not phenomenal quality at all (p.394)

According to type materialism, the explanation is that the neuro-scientific state (or quality) in question is identical with the phenomenal state (or quality). According to psychofunctionalism, the explanation is that the phenomenal state is identical with a deep functional role state, and the neuro-scientific state realizes that state.

To be sure, these would-be solutions face a host of familiar difficulties, including the appearance of an explanatory gap. But to belabor the point: the difficulty posed by the issue of Commander Data's consciousness is supposed to be different from the familiar difficulties faced by physicalist solutions to the hard problem. I shall thus argue for a hypothetical claim: *if* we have a conception of how the familiar problems that beset type materialist and psychofunctional solutions to the hard problem can be adequately handled without compromising phenomenal realism, then we have a conception of how we could have rational grounds for denying that Commander Data is conscious that involves no compromise of phenomenal realism. To be sure, that is a big "if." But Block acknowledges that we have "a glimmer of a conception" of how the familiar problems for physicalist solutions to the hard problem can be adequately addressed. He ought, then, I claim, to acknowledge that we have a glimmer of a conception of how we could have rational grounds for denying that Commander Data is conscious.

The truth of type materialism depends, of course, on the truth of the following thesis:

*Neuro-Scientific Nomological Correlation Thesis*. For every state of phenomenal consciousness, C, there is some neuro-scientific state, N, such that it is a law that a being is in C if and only if that being is in N.

And the truth of psychofunctionalism depends on the truth of this thesis:

Deep Functional Nomological Correlation Thesis. For every state of phenomenal consciousness, C, there is some deep functional role state, F, such that it is a law that a being is in C if and only if that being is in F.

We do not now know whether either of these nomological correlations theses is true. Of course, if neither is true, then neither type materialism nor psychofunctionalism is true. It could turn out, however, that the neuro-scientific correlation thesis is false, and that the deep functional correlation thesis is true. If so, then type materialism, at least, is false. Or it could turn out that the deep functional correlation thesis is false, and that the neuro-scientific correlation thesis is true. If so, then psychofunctionalism, at least, is false. Or it could turn out that both correlation theses are true. That both correlation theses are true is logically compatible with type materialism and psychofunctionalism.<sup>37</sup>

Type materialism and psychofunctionalism can differ over whether a hypothetical robot (or alien being) is conscious. Suppose that a psychofunctionalist theory identifies a state of consciousness, C, with a deep functional state, F, and that it is possible for a robot to be in F without being in any of our neuro-scientific states, including the abstract neuro-computational ones. Thus, to be in F would be to be in some state with a certain causal role, R, where R does not include any of our neuro-scientific state-types, or any state types that necessitate them. Then, psychofunctionalism will entail that the robot is conscious, while type materialism will entail that it is not conscious. This dispute over robot consciousness need not concern us here, however. For, by stipulation, Commander Data fails to share any neuro-scientific or deep functional role state with us. The case of Commander Data is thus easier. If we have a conception of how we could have rational grounds for believing either type materialism *or* psychofunctionalism without compromising phenomenal realism, then we have a conception of how we could have rational grounds for believing that Commander Data is not conscious that involves no compromise of phenomenal realism.

The only (relevant) way to have rational grounds for believing type materialism is by inferring the thesis as the best explanation of the neuro-scientific correlation thesis; and similarly for psychfunctionalism and the deep functional correlation thesis. Because the differences between type materialism and psychofunctionalism do not matter for present purposes, and for simplicity of exposition, I shall speak ambiguously of 'the correlation thesis', and ambiguously of 'physical-phenomenal identity claims'. I shall do this, even though were both correlation theses true, type materialism and psychofunctionalism would offer different explanations as to why.

<sup>&</sup>lt;sup>37</sup> I should mention that as concerns each of the correlation theses, there will be issues that arise concerning semantic indeterminacy due to vagueness. I cannot address such issues here. See McGee and McLaughlin 2000.

Suppose, then, that we had rational grounds for believing that the correlation thesis held true for the restricted domain of human beings, members of neighboring species that we intuitively regard as "undoubtedly conscious," and members of some more distant species for which there had been some prior question about their consciousness. Block readily acknowledges that we could have such rational grounds (p.000). However, he holds that Commander Data and his like would pose an obstacle to generalizing the correlation thesis to all beings. For were we to so generalize it, we would of course be committed not only to denying that, say, gold fish are conscious, but also that Commander Data and his like are. In reply, I claim that Commander Data would pose no such obstacle, provided that a certain explanatory condition is met. I claim that the nomological possibility of Commander Data would not be an obstacle to generalizing the correlation thesis to all beings, provided that we could infer either type materialism or psychofunctionalism as the best explanation of the thesis.

The problem of robot consciousness now rears its head in the following objection. The physical-phenomenal identity statements would provide us with rational grounds for believing that Commander Data is not conscious only if we had rational grounds for believing them. And we could have rational grounds for believing them only if we could infer them as the best explanation of the generalized correlation thesis. But we could be in a position to do that only if we had rational grounds for believing the generalized correlation thesis. How could we have rational grounds for believing the generalized correlation thesis unless we already had rational grounds for believing that Commander Data is not conscious? While I have argued that we have no (undefeated) reason whatsoever to maintain that Commander Data is conscious, the issue remains whether we have rational grounds for believing that Commander Data is not conscious. And it seems that we would have rational grounds for believing the correlation thesis only if we had rational grounds for believing that Commander Data is not conscious. So we are back to square one.

In reply, it is indeed true that we would have rational grounds for believing the physical-phenomenal identity claims only if we had rational grounds for believing the correlation thesis (generalized), and that we would have rational grounds for believing that thesis only if we had rational grounds for believing that Commander Data is not conscious. But it would be a mistake to think that the belief that Commander Data is not conscious is *epistemically prior* to the belief in the correlation thesis, and that it, in turn, is epistemically prior to the belief in the identity thesis. That is not so. Notice, first of all, that our rational grounds for believing something can include the explanatory coherence of the belief. Explanatory coherence is not hierarchical. Beliefs can be mutually supporting in virtue of their explanatory coherence.<sup>38</sup> Here is a simple example. Because the weather forecast that I saw last night predicted rain during the early hours, I awake in the morning with the belief that it rained last night. My belief that it rained last night provides me with a reason to believe that my grass is wet. I then look out the window expecting, for that reason, to see wet grass. The grass indeed looks wet. My belief that the grass is wet offers further support to my belief that it rained last night. There is

This idea is defended in Sosa (1997), and in Sosa forthcoming.

mutual support through explanatory coherence.<sup>39</sup> I claim that the relationship that the belief in a property or kind identity bears to the belief in their correlation is typically one of mutual explanatory coherence; and that these beliefs can join to warrant belief in some instances of the correlation that might otherwise be undecided.

Consider a paradigm case of inferring identity as the best explanation of a correlation, namely the case of inferring that water =  $H_2O$  as the best explanation of the correlation of water and  $H_2O$ . As Block himself correctly points out in his discussion of the epistemology of theoretical identity,

Deciding that water and  $H_2O$  are spatiotemporally coincident is part of the same package as deciding they are one and the same. For example, the air above a glass of water buzzes with bits of water in constant exchange with water in the atmosphere, a fact that we can acknowledge only if we are willing to suppose that those  $H_2O$  molecules *are* bits of water. The claim that water is  $H_2O$  and that water and  $H_2O$  are spatiotemporally coincident stand or fall together as parts of one explanatory package (p.410).

That is right. Consider, then, the following objection to the view that water =  $H_2O$ . We could determine whether water and H<sub>2</sub>O are identical only if we could determine that they are (of nomological necessity) spatiotemporally coincident. We know by sophisticated tests that individual H<sub>2</sub>O molecules will escape from boiling water. Water and H<sub>2</sub>O are spatiotemporally coincident only if there is water where each such H<sub>2</sub>O molecule is. At best, folk chemistry leaves the issue open whether there is water where such H<sub>2</sub>O molecules are. Thus, until we can determine whether there is water wherever each such H<sub>2</sub>O molecule is, we cannot determine whether water is H<sub>2</sub>O. Now as I noted, explanatory coherence does not involve the sort of epistemic priority in question. That water and  $H_2O$  are correlated is part of our reason for holding that water =  $H_2O$ . (I say 'part of our reason' since it is not the whole of it: there are as well reasons of theoretical simplicity and overall coherence.) And our belief that water =  $H_2O$ , one we infer as the best explanation of the correlation, contributes to providing us with rational grounds for the belief that they are correlated. (I say 'contributes' because observation makes a crucial contribution too.) These beliefs, in turn, provide rational grounds for the belief that H<sub>2</sub>O molecules that escape boiling water are water. Likewise, the physicalphenomenal identity claims, the physical-phenomenal correlations, and the negative judgment on Commander Data, would stand or fall together as parts of one explanatory package. There are, to be sure, striking dissimilarities between our case of interest and the case of water. But the relevant similarity holds: there would be mutual support through explanatory coherence.

An essential part of my story about how we could have rational grounds for believing that Commander Data lacks consciousness is that we could be in a position to infer identity as the best explanation of the correlation thesis. But there are, of course, a host of familiar problems that beset attempts to so infer, which have no counterpart in the case of "Water =  $H_2O$ ." I shall formulate the leading problems claimed to be of this sort

Examples of this sort are examined in considerable detail in Sosa forthcoming.

as challenges for type materialism and psychofunctionalism, and follow them with brief responses.

First, there is what I shall call *the Huxley-McGinn Challenge*: "How is it that anything so remarkable as a state of consciousness comes about a result of irritating nervous tissue, is just as unaccountable as the appearance of Djin when Aladdin rubbed his lamp" (Huxley 1886, p.193). "How is it possible for conscious states to depend upon brain states? How can technicolour phenomenology arise from soggy grey matter?" (McGinn 1989, p. 349) Indeed how could a physical state give rise to a conscious state? What could the mechanism be? *Reply*. There is no mechanism. The relevant physical states do not give rise to conscious states. Rather, they are conscious states.

Second, there is what I shall call the Horgan-Tienson's Challenge: 42 we have direct introspective access to our states of phenomenal consciousness, and thus our access to them is not mediated by contingent modes of presentation of them or aspects of them. But we do not introspect our states of consciousness either as neuro-scientific states or as psycho-functional role states. How, then, given these facts about introspection, could a state of consciousness be either a neuro-scientific or a deep functional role state? Reply. Our introspective access to our states of consciousness is indeed experientially direct: it is unmediated by other experiences or partial aspects of the conscious states. (That is why introspection is not a kind of perception.) It is also true that introspection does not reveal a conscious state either as a neuro-scientific state or as a deep functional role state. But it is not the case that introspection reveals a conscious state to be otherwise than a neuroscientific state or a deep functional role state. Introspection-as involves the application of concepts; and we cannot deploy neuro-scientific concepts or psychofunctional concepts directly in introspection. Nonetheless, the phenomenal concepts that we can directly employ in introspection answer either to neuro-scientific properties or to deep functional properties.<sup>43</sup>

Third, there is what I shall call *the Nagel Challenge*:<sup>44</sup> states of consciousness are subjective states, while neuro-scientific states and psychofunctional states are objective states. How could a subjective state be an objective state? How could an objective state be a subjective one? *Reply*.<sup>45</sup> The subjective/objective distinction is, in the first instance, an epistemic distinction; a state or property is only subjective or objective under a concept. There is no a priori link between subjective concepts and objective ones. But a state type can be subjective under one concept, while being objective under another.<sup>46</sup>

11

<sup>&</sup>lt;sup>40</sup> Block cites this quotation.

<sup>&</sup>lt;sup>41</sup> See, e.g., Papineua 2002.

<sup>42</sup> Horgan and Tienson 2001.

<sup>&</sup>lt;sup>43</sup> For further discussion, see McLaughlin 2001 and 2003.

<sup>44</sup> Nagel 1974.

<sup>45</sup> See, e.g., Loar 1990/7; Sturgeon 1994; Perry 2001; and McLaughlin 2003.

<sup>&</sup>lt;sup>46</sup> Block raises this challenge and discusses this sort of response to it with some approval. His approval is, however, qualified. He says: "As Nagel noted, we have no idea how there could be causal chains from an objective concept and a subjective concept leading back to the same phenomenon in the world" (p.398). But we know how there could be causal chains from an objective concept and a subjective concept leading back to the same phenomenon in the world: that could happen if the properties that the concepts refer to are

Fourth, there is what I shall call the Kripke Challenge: 47 the physical-phenomenal identities in question would be are necessary. But since they would also be only a posteriori knowable, in asserting them the type materialist and psychofunctionalist incur the dialectical obligation of explaining away their appearance of contingency. Moreover, Kripke's model for explaining away the appearance of contingency of only a posteriori knowable identity statements such as "Water =  $H_2O$ " won't serve this purpose in the cases in question since epistemic counterparts of states of consciousness are states of consciousness. (Thus, for instance, any epistemic counterpart of pain is pain.) Rely. While we indeed cannot employ Kripke's model for the reason cited, we can construct an alternative model for explaining away the appearance of contingency of identity statements that are only a posteriori knowable, a model that is applicable to physicalphenomenal identity claims. One such model that has much to commend it is Christopher Hill's (1998) image-splicing and concept-splicing model.<sup>48</sup>

Finally, there is what I shall call the Levine-Chalmers-Jackson Challenge: 49 these physical-phenomenal identity statements are "gappy" in that they are not a priori entailed by the totality of micro-physical facts, the fact that that totality is the totality of microphysical facts, and the assumption that our world is a minimal physical duplicate of itself.<sup>50</sup> But Kripke's paradigms of a posteriori statements, statements such as "Water = H<sub>2</sub>O" are so a priori deducible. Thus, there is an explanatory gap as concerns such physical-phenomenal identity statements, but not as concerns Kripke's paradigms. The challenge, then, is to show how the physical-phenomenal identity claims could be justified, given this explanatory gap.

There are two (compatible) lines of response to this challenge that one finds in the literature. Response 1. Not even Kripke's paradigm, "Water = H<sub>2</sub>O," is a priori deducible from such facts; and so, the psycho-physical identity claims asserted by type materialists and psychofunctionalists are not special in the respect in question.<sup>51</sup> Moreover, they are justified on the grounds that they provide the best explanations of physical-phenomenal

identical. To be sure, we need a theory of conceptual reference that permits a subjective concept to have the same referent as an objective one. But there is no obstacle to it being a causal theory (as opposed to an a priori causal-descriptive theory) of some sort, for such a theory will appeal to underlying causal referencedetermining mechanisms to which we lack any a priori or introspective access. Block adds: "We are in something like the position of pre-Einsteinians who had no way of understanding how a concept of mass and a concept of energy could pick out the same thing" (p.398). I do not think that the development of new concepts will help "close the explanatory gap." For subjective concepts and objective concepts are different in such a way that we could not develop concepts that provided an a priori bridge between them. It is better to explain why there could be no such bridge by appealing to the different roles of subjective and objective concepts in our cognitive economy.

31

Kripke 1980.

See also Hill 1998; Hill and McLaughlin 1999; and McLaughlin 2001, 2003. Hill's model is inspired by some remarks in Nagel 1974; and it has close affinities with Papineau's (2002) explanation of the appearance of an explanatory gap between the mental and the physical. See also Lycan's (1987) discussion of "the stereoscopic fallacy."

<sup>&</sup>lt;sup>49</sup> Levine 1983, 2000; Jackson 1993, 1998; and Chalmers 1996.

A physical duplicate of the actual world is a world exactly like the actual world in every physical respect. A minimal physical duplicate of the actual world is a world that contains only whatever is metaphysically required to be a physical duplicate of the actual world. (See Jackson 1993; 1998.) <sup>51</sup> Block-Stalnaker 1999.

nomological correlations. *Response 2*. Even if the physical-phenomenal identity statements in question are special in this way, that would be explained by a fact entailed by phenomenal realism, namely that phenomenal concepts lack either strong or weak a priori functional analyses. They lack strong functional analyses because they are not definable by physical and topic neutral concepts. And they lack weak functional analyses because there are no contingent a priori reference-fixing conditions for them. In this last respect, they differ from such concepts as the concept of water, *if* that concept indeed has a contingent a priori reference-fixing condition. Phenomenal concepts differ from physical concepts markedly in their conceptual roles. To note repeat difference: we can apply the former, but not the latter, directly in introspection. The special features of phenomenal concepts explain why there has been a mind/body problem for over three hundred years. But what are special are phenomenal concepts. It does not follow that the properties that answer to them are special.<sup>52</sup>

There are some further familiar challenges that could be considered; and there are further replies. Moreover, much more can be said, and indeed has been said elsewhere, by way of development of the above replies. But the pattern should now have emerged. It has four steps. First, acknowledge the epistemic claim made by the dualist (e.g., there are no a priori links between physical concepts and phenomenal concepts, introspection does not reveal experiences to be neuro-scientific or deep functional role states, it appears that some things can be known only subjectively and that some can be known objectively, physical-phenomenal identity claims have an appearance of contingency that must be explained away, psychophysical identities are not a priori deducible from physical facts). Second, explain why it is true by appeal to the differences in the conceptual roles of phenomenal concepts and physical(/functional) concepts. Third, point out that the explanation does not entail that conscious states are not neuro-scientific or deep functional states. Fourth, argue, on grounds of overall coherence and theoretical simplicity, that phenomenal concepts and physical concepts answer to the same properties.

Now those who reject naturalism will no doubt fail to be satisfied by what I have said. But my main audience here is naturalist-phenomenal realists. My claim, you will recall, is a conditional one: *if* we have a conception of how the familiar difficulties that beset type a type materialist or psychofunctionalist solution to the hard problem can be adequately addressed without compromising phenomenal realism, then we have a conception of how we could have rational grounds for denying that Commander Data is conscious that involves no compromise of phenomenal realism.

What explains, then, lingering doubt about whether Commander Data is conscious? We do not of course now know whether either type materialism or psychofunctionalism is true. First, while I think that we have more than a mere glimmer of a conception of how to resolve the familiar issues surrounding physicalist solutions to the hard problem, the issues have not yet been resolved. Second, while it is an

See, e.g., Loar 1990/7; Hill 1998; Hill and McLaughlin 1999; and McLaughlin 2001 and 2003.

Fig. 1990/7, Hill 1998, Hill and McLaughlin 1999, and McLaughlin 2001 and 2003.

See, e.g., Loai 1990/7, Hill 1998, Hill and McLaughlin 1999, and McLaughlin 2001 and 2003.

I have not mentioned Jackson's (1982) case of Mary since I think that, with an added bell and whistle, it can be handled by the response to the Nagel objection. See McLaughlin 2003.

increasingly reasonable speculation, we do not now know whether states of phenomenal consciousness are nomologically correlated with either neuro-scientific states or psychofunctional states even in humans and members of neighboring species. I do not, however, think that these considerations play a major role in explaining lingering doubt about whether Commander Data is conscious.

I think that part of the explanation for lingering doubt is that it is easy to slip into a skeptical frame of mind. Given phenomenal realism, there is room for philosophical doubt about whether Data is conscious. The same, however, is true of a laptop or a rock. Might not laptops and rocks have a kind of phenomenal buzz? How does one know that they do not? On the other hand, how does one know that one's fellow human beings are conscious? Indeed, how does one know that one did not just become conscious for the very first time a few moments ago? But the issue, you will recall, is whether we could have what passes in the practice of special sciences as rational grounds for belief that Commander Data is not conscious. The issue of Commander Data's consciousness is not supposed to collapse into a case of the skeptical problem of other minds.

There are further sources of lingering doubt. One might think that I have underestimated the epistemic value of superficial functional isomorphism as a reason for attributing phenomenal consciousness. One's intuitions might not accord with mine about the cases discussed earlier (robot rabbits, and the like). Of course, often intuitions can be changed by careful reflection. Nevertheless, I take intuitions very seriously since we will not believe that we have an explanation of a phenomenon unless the would-be explanation accords at least reasonably well with our considered judgments. There is, however, more that I can say by way of trying to show what is behind the intuition of the weightiness of superficial functional isomorphism as a reason for attributing phenomenal consciousness.

I think that a large part of what is behind this intuition is that when we contemplate robots like Commander Data, we realize that if we had to interact extensively with them, we would have to treat them as if they were phenomenally conscious; for that would be the only way that we could predict and make sense of their behavior. Notice, however, that the same is true of homunculi heads and partial physical overlappers. In order to interact effectively with them, we would have to so treat them. In acknowledging this, we need not concede that they are phenomenal consciousness. The point, rather, is that to deal with them effectively, we would have to take what we might call 'the sentient stance' toward them. We can, however, distinguish the claim that it would be necessary to use the vocabulary of phenomenal consciousness to predict and make sense of their behavior from the claim that we are justified in applying phenomenal predicates (e.g., 'is in pain') in their normal senses to them. Rather, the terms would be used in an analogical sense.<sup>54</sup>

David Chalmers (1996) has claimed that there is a distinction between the psychological or (to use more a accurate term) purely functional use of phenomenal terms (such as 'pain') and their phenomenal use. If he were right that there is a purely

-

<sup>&</sup>lt;sup>54</sup> See Hill 1991. Chapter 9.

functional, deflationary use of these terms, no doubt that would explain our intuition concerning the weightiness of superficial functional considerations. There would be superficial functional conditions that are analytically sufficient for being in pain, having an itch, etc., *in the purely functional senses in question*. Superficial functional isomorphism would analytically guarantee phenomenal consciousness in a purely functional sense of 'phenomenal consciousness'. In thinking this weight carries over to attributions of phenomenal consciousness in the relevant sense, however, we would be confusing functional pain (and the like) with phenomenal pain (and the like).

I do not think that there are two such established uses of terms like 'pain', 'itch', etc.. But I think that there could come to be. Regular intercourse between us and a community of androids like Data might very well lead to a change in the use of terms like 'pain'; the terms might come to have a regular, ordinary purely functional use that is entirely distinct from their phenomenal use. Should that come to pass—or if Chalmers right that it is now in fact the case—, then applying 'pain', 'itch', and the like to states of Data would not simply be a matter of stance-taking. Such predicates would literally apply to Data's states; but they would apply to Data's states only in their purely functional sense, not in their phenomenal sense. I think that our tacit recognition of this possibility contributes to our sense of the weightiness of superficial functional organization as a reason. But we must not confuse purely functional uses of these terms with their phenomenal uses.

I think that another concern behind the issue of Commander Data's consciousness is that Data appears to have beliefs about his own mental states. Indeed, as I mentioned in part 1, we might have been asked to suppose that Commander Data has such beliefs when we were asked to suppose that Data is a superficial functional isomorph of us (normal, adult human beings). In any case, to belabor yet another point: the issue is whether Commander Data has phenomenal consciousness, not whether Data propositional attitudes or even higher-order propositional attitudes.

Yet a further concern behind the issue of Commander Data's consciousness is the sense that Data is a being worthy of moral respect. Indeed at one point in "The Harder Problem of Consciousness," Block alludes to an episode of "Star Trek: The New Generation" entitled 'The Measure of the Man', in which there is a trial to determine whether the Star Fleet has the right to take Data apart, despite his (apparent) desire not to be taken apart. Block seems to suggest that anyone who watches this episode will find it hard, or at least ought to find it hard, to resist the feeling that Commander Data should not be taken apart. If Commander Data is truly devoid of phenomenal consciousness, could this sentiment be warranted? Or are we just misled by Commander Data's human-like appearance—Data's face, eyes, etc.—and Data's human-like behavior?

Suppose that despite lacking phenomenal consciousness, Commander Data in fact has higher-order beliefs, desires, intentions, and the like. If Data is such a higher-order intentional being, then Data is a *moral agent*, and thus an appropriate subject of praise or blame. The question remains, however, whether Data is a *moral patient*: do we have moral obligations to Data? Even if Data is an appropriate subject or praise or blame, do

we have even a *prima facie* obligation to respect Data's wishes, and so not take Data apart? It might be claimed that we do only if it would harm Data; but since Data lacks phenomenal consciousness, Data cannot suffer, and so cannot be harmed. 55 Now in a purely functional sense of 'suffer', Data can suffer; in a purely functional sense of 'pain', Data can feel pain. Data is able to detect bodily damage, desires to avoid bodily damage, and so. Of course, it might well be claimed in response that this fails to suffice for harm in the sense in question. If, however, having phenomenal consciousness is necessary for suffering harm in the sense in question, then an appropriate counter-response is that even if we cannot *harm* Data, we can *wrong* Data. Given that Data has higher-order intentional states, Data can pursue a life-plan; Data is an end-in-itself. We would disrespect Data by blocking him from the achievement of his life goals. These considerations, I believe, would warrant the view that we have moral obligations to Data.

Consider (philosophical) zombies. They are exactly like us in every physical respect, but are devoid of consciousness. Phenomenal realists of all stripes hold that such zombies are coherently conceivable. (Of course, naturalist-phenomenal realists nevertheless deny that they are possible, while dualist-phenomenal realists maintain that they are possible.) It seems deeply intuitive that such zombies would have moral obligations to one another, and could wrong one another. The moral: zombies would be people too. The consequence: theories of morality that make essential appeal to phenomenal consciousness are not a priori true.

Of course, each year we see robots of increasing complexity. Robots like Commander Data, however, are not now on the horizon. And while they are not beyond the bounds of our imagination, they may well prove to be beyond our capacities of ingenuity, or even beyond the bounds permitted by laws of nature. If, however, as science fiction writers claim to foresee, we someday face the serious prospect of robots like Commander Data—robots that are superficial functional isomorphs of us—the question of whether they are deserving of moral respect will be more than academic: it will be a pressing social and moral issue. We can now, however, reflect on the question without the burden of immediacy of social and moral consequence. I contend that we have a hypothetical conception of how we could have rational grounds for denying that they are phenomenally conscious. Even if they would indeed lack phenomenal consciousness, however, if they have higher-order intentional states, then it would be wrong to treat them as any less than full members of our moral community. <sup>56</sup>

#### References:

Of course, it might do harm to our moral characters to so treat Data, given the closeness of Data to ourselves in our psychological similarity space. And think of the effect on the moral education of children. But am I concerned here with the issue of harm and wrong to Data.

<sup>&</sup>lt;sup>56</sup> This paper began life as a comment on Block's "The Harder Problem of Consciousness," presented at the December 2001 SOFIA conference. I wish to thank the audience there and Block himself for comments. I also wish to think the audience at the University of Carleton, where an early draft of this paper was presented in 2002. Thanks are also owed to John Hawthorne, Peter Klein, and Ernest Sosa for very helpful comments on the paper. Finally, I owe a deep debt to the pioneering work of Christopher Hill (1992) on type materialism.

- Block, N. 1980. 'Troubles with Functionalism' in his <u>Readings in the Philosophy of Psychology, Vol.I.</u> Cambridge, Mass.: Harvard University Press, pp.268-305.
- Block, N. and Stalnaker, R. 1999. 'Conceptual Analysis and the Explanatory Gap', *The Philosophical Review*, CVIII, 1-46.
- Block, N. 2002. 'The Harder Problem of Consciousness'. *The Journal of Philosophy*, Vol.XCIX, 8: 391-425.
- Block, N. 2003. 'The Harder Problem of Consciousness' www.nyu.edu/gsas/dept/philo/faculty/block/papers/harder.htm.
- Chalmers, D. 1995. 'Facing Up to the Problem of Consciousness', *Journal of Consciousness Studies* 4:3-46.
  - Chalmers, D. 1996. The Conscious Mind. New York: Oxford University Press.
- Chalmers, D. 2002. 'Consciousness and its Place in Nature' in ed. Chalmers, D. <u>Philosophy of Mind: Classical and Contemporary Readings</u>. New York: Oxford University Press, pp.247-72.
- Dennett, D.1978. 'Why You Can't Make a Computer that Feels Pain' in ed. Dennett, D. <u>Brainstorms</u>. Montgomery. Vt.: Bradford Books, 190-232.
- Erwin, E. 1970. <u>The Concept of Meaninglessness</u>. Baltimore: The John Hopkins Press.
- Field, H. 1996. 'The Aprioricity of Logic', *Proceedings of the Aristotelian Society* 96:325-76.
- Gunderson, K. 1971. <u>Mentality and Machines</u>. Minneapolis: University of Minnesota Press.
- Hill, C. 1991. <u>Sensations: A Defense of Type Materialism</u>. Cambridge: Cambridge University Press.
- Hill, C. 1997. 'Imaginability, Conceivability, Possibility, and the Mind-Body Problem', *Philosophical Studies* 87:61-85.
- Hill, C., and McLaughlin, B. 1999. 'There are Fewer Things Than are Dreamt of in Chalmers' Philosophy.' *Philosophy and Phenomenological Research* 2: 445-54.
- Horgan, T. and Tienson, J. 2001. 'New Wave Materialism', in B. Loewer and C. Gillett (eds.) <u>Physicalism and Its Discontents</u>. Cambridge: Cambridge University Press.

- Hurvich, L.M. 1981. Color Vision. Sunderland, Mass.: Sinauer.
- Huxley, T.H. 1896. <u>Lessons in Elementary Physiology</u>. London: Macmillan.
- Jackson, F. 1982. 'Epiphenomenal Qualia', Philosophical Quarterly 32, 127-36.
- Jackson, F. 1993. 'Armchair Metaphysics' in J. O'Leary-Hawthorne and M. Michaels (eds.) <u>From Concepts to Metaphysics</u>. Oxford: Oxford University Press.
- Jackson, F. 1998. <u>From Metaphysics to Ethics: A Defense of Conceptual Analysis</u>. Oxford: Oxford University Press.
- Kripke, S. 1980. <u>Naming and Necessity</u>. Cambridge, Mass.: Harvard University Press.
- Levine, J. 1983. 'Materialism and Qualia: The Explanatory Gap', *Pacific Philosophical Quarterly*, LXIV, 354-61.
  - Levine, J. 2001. Purple Haze. Oxford: Oxford University Press.
- Lewis, D. 1966. 'An Argument for the Identity Theory', *Journal of Philosophy*, LXIII, 17-25.
- Lewis, D. 1972. 'Psychophysical and Theoretical Identifications', *Australasian Journal of Philosophy*, 50, 249-58.
- Loar, B. 1990/7. 'Phenomenal States', in (eds.) Block, N., Flanagan, O., and Guzeldere, G. <u>The Nature of Consciousness</u>. Cambridge, Mass.: MIT Press.
- Long, D.C. 1994. 'Why Machines Can Neither Think Nor Feel', in ed. Jamison, D. <u>Language, Mind, and Art: Essays In Appreciation and Analysis</u>, In Honor of Paul Ziff. Dordrecht: Kluwer Academic Publishers), 101-119.
  - Lycan, W. 1987. Consciousness. Cambridge, Mass.: MIT Press.
  - Lycan, W. 1996. Consciousness and Experience. Cambridge, Mass.: MIT Press.
- MacKay, D.M. 1952. 'Mentality in Machines', *Aristotelian Society Supplement*, XXVI, 61-68.
- MacKay, D.M. 1956. 'The Epistemological Problem for Automata', <u>Automata Studies</u>. Princeton: Princeton University Press, pp.235-51.
  - McGinn, C. 1989. 'Can We Solve the Mind-Body Problem', Mind 98:349-66.

- McLaughlin, B. 2001. 'In Defense of New Wave Materialism', in eds. Loewer, B. and Gillett, K. <u>Physicalism and Its Discontents</u>. Cambridge: Cambridge University Press.
- McLaughlin, B. 2003. 'Color, Consciousness, and Color Consciousness' in eds. Smith, Q, and Jokic, A. <u>Consciousness: New Philosophical Perspectives</u>. Oxford: Oxford University Press.
- Nagel, T. 1974. 'What Is It Like to Be a Bat?', *The Philosophical Review*, LXXXIII: 435-50.
- Papineau, D. 2002. <u>Thinking About Consciousness</u>. Oxford: Oxford University Press.
- Perry, J. 2001. <u>Knowledge, Possibility and Consciousness</u>. Cambridge: MIT Press.
- Pollack, J. 1974. <u>Knowledge and Justification</u>. Princeton: Princeton University Press.
- Pollack, J. 1986. <u>Contemporary Theories of Knowledge</u>. Totowa, NJ: Roman and Littlefield.
- Putnam, H. 1964. 'Robots: Machines or Artificially Created Life? *Journal of Philosophy*, LXI, 668-91.
  - Scriven, M. 1953. 'The Mechanical Concept of Mind', Mind, LXII, 230-240.
  - Searle, J. 1992. The Rediscovery of the Mind. Cambridge: MIT Press.
- Sider, T. 2003. 'Maximality and Mircophysical Supervenience', *Philosophy and Phenomenological Research* LXVI. 1: 139-149.
- Sosa, E. 1997. 'Reflective Knowledge in the Best Circles', *Journal of Philosophy*, XCIV, 8:410-30.
  - Sosa, E. Forthcoming. 'Virtual Epistemology'.
  - Strawson, G. 1994. Mental Reality. Cambridge: MIT Press.
- Sturgeon, S. 1994. 'The Epistemic View of Subjectivity', *Journal of Philosophy*, XCL, 5:221-35.
- Sober, E. 2000. 'Evolution and the Problem of Other Minds' *Journal of Philosophy*, XCVII, 71 365-86.

Turing, A.M. 1950. 'Computing Machinery and Intelligence', *Mind*, LIX, 433-466.

Ziff, P. 1959. 'The Feelings of Robots', Analysis, XIX, 241-249.