

THE PROPERTY DUALISM ARGUMENT

I. THE EXPLANATION OF A POSTERIORI IDENTITIES

Suppose that Smith's pain at t is identical with Smith's C-fiber firing at t . More specifically, suppose that the token state or event that has the property at t of being Smith's only pain also has the property at t of being his only C-fiber firing.¹ Assume that this identity is an empirical fact, discoverable only a posteriori. Then as used by Smith at t , the expressions 'my pain' and 'my c-fiber firing' will be coreferential, but will not (in any intuitive sense) mean the same thing. We must, therefore, explain how a mentalistic expression and a physicalistic expression could refer to the same event while at the same time satisfying the following desiderata.

(1) We explain how an identity such as 'My pain is identical with my C-fiber firing' could be a posteriori; i.e.,

(1') we explain the difference in the cognitive significance of such expressions as 'my pain' and 'my C-fiber firing'; i.e.,

(1'') we explain how a subject could be fully rational in believing what he or she would express by saying such things as "I am in pain" and also in believing what he or she would express by saying things such as "It is not the case that my C-fibers are firing," and, more generally, anything of the form 'It is not the case that I am in a state of kind D' where 'D' stands in for a physical and/or functional description.

(2) We satisfy Frege's constraint in what I shall call its ordinary version (OVFC): If x believes y to be F and also believes y not to be F, then (to the extent that x is rational) there must be distinct representational modes of presentation \underline{m} and \underline{m}' such that x believes y to be F under \underline{m} and disbelieves y to be F under \underline{m}' .²

(3) We satisfy Frege's constraint in what I shall call its strong version (SVFC): We satisfy (1) and (2) by satisfying the following conditions.

- (a) The representational modes of presentation referred to in (2) provide a rational justification of the subject's beliefs, intentions, and actions.
- (b) The justification is available to the subject at the personal level.
- (c) The justification takes the form of a characterization of the way the world presents itself to the subject or the way the world is given from the subject's point of view.

As it is used in the literature, 'mode of presentation' is ambiguous between something on the side of language and content, such as a description, concept, or some form of nonconceptual content,³ and something on the side of the world, such as a property.⁴ Thus in the conditions above and in what follows I distinguish between representational modes of presentation (on the side of content) and nonrepresentational modes of presentation (on the side of the world).

What the strong version of Frege's constraint adds to versions such as OVFC is an explicit acknowledgment of the justificatory role that modes of presentation in both senses are required to play. Modes of presentation are not postulated simply to explain behavioral and/or functional dispositions--for example, a subject's disposition to produce the sound associated with 'yes' in response to 'Are you in pain?' and that associated with 'no' in response to 'Are your C-fibers firing?' For this role a difference in the causal chains connecting the subject's pain to tokens of the word 'pain' on the one hand and to tokens of 'C-fiber firing' on the other (and a resulting difference in the causal role of the two terms in the subject's functional economy) would suffice. But such a difference in the causal chains and the functional roles associated with different linguistic expressions is one to which the subject need not have access. Thus, in the absence of some personal-level manifestation, such a causal difference could not play the justificatory role for which modes of presentation are slated

Why, though, should we hold out for a justificatory as opposed to explanatory role for modes of presentation? The first reason is that conditions (3a)-(3c) are already implicit in conditions (1) and (2). Thus a justification of the latter conditions will provide a justification of the former as well.

Condition (1) is a given in this context and is accepted on all sides; (1') and (1'') are intended as elucidations of (1). Furthermore, (1') and (1'') (and thus (1)) will be accepted by anyone who takes modes of presentation seriously and thus by anyone who takes Frege's constraint seriously in its ordinary

version. Moreover, as I shall argue below, condition (2), Frege's constraint in its ordinary version, should be accepted by anyone who is serious about belief ascription. The same argument, which appeals to the constitutive role of charity and rationality in the ascription of intentional states, is sufficient to justify (3a) as well. And it is clear, as we have seen, that such rational justification must take place at the personal level (as (3b) requires).

To support the claim that the kind of rational justification in question proceeds by characterizing the way the world presents itself to the subject (i.e., (3c)), we need simply note that beliefs (and such other intentional entities as intentions and actions) are justified by appeal to other beliefs and intentional states and other representational contents—including, possibly, nonconceptual contents. Hence they are justified by reference to the way the world presents itself or the way in which it is given. But we have already seen that rational justification is a matter of what is available to the subject at the personal level. Hence rational justification is a matter of the way the world presents itself to the subject or the way it is given from the subject's point of view.

Condition (3), SVFC, has a second, related source of support. To suppose that differences in the modes of presentation required by Frege's constraint could consist merely in causal differences unavailable to the subject at the personal level is to adopt a position that I shall call local eliminativism. Just as the eliminativist regarding intentionality eschews talk of content altogether in favor of an explanation of behavior in terms drawn from the natural sciences, so the local eliminativist (regarding intentionality) eschews such talk in what we might call the Fregean contexts (i.e., the contexts in which SVFC would require the postulation of representational modes of presentation available to the subject in question). But local eliminativism is not a tenable position. The point of ascribing intentional content is to characterize the world as it presents itself to the subject, thereby providing a rational justification of the subject's beliefs, intentions, and actions. When we do so I shall say that we rationalize those intentional states (and to that extent their subject). In fact, rationality is constitutive of the project of intentional ascription, as is evidenced by the constitutive role of the principle of charity in radical interpretation.⁵ Indeed, it would be a mistake to think of our commitment to Frege's constraint and to the principle of charity as independent. They are more appropriately seen as different manifestations of our fundamental commitment to the rationality of the subjects of intentional states.⁶ Local eliminativism should be rejected

in favor of full-blown eliminativism by those who are skeptical about intentional ascription altogether. Those who are not should hold out for the higher standard that SVFC entails.

II. THE ARGUMENT FOR PROPERTY DUALISM

I shall now outline the property dualism argument,⁷ some points of which will receive further elaboration in subsequent sections. Notice first that although the explanations required by conditions (1)-(3) call for the rational justification of the beliefs, intentions, and actions ascribed to the subject, this is not to say that such contents play no explanatory role. The contents ascribed to such subjects must explain the subjects' access to the referents of their expressions. For example, to provide the explanations called for by (1) regarding the identity statement 'My pain is identical with my C-fiber firing' while satisfying OVFC, the two expressions involved must pick out the objects they do under distinct representational modes of presentation. And corresponding to these representational modes of presentation there must be properties (i.e., nonrepresentational modes of presentation) in virtue of which they pick out the object in question. Moreover, the representational and nonrepresentational modes of presentation must be appropriately related; indeed the relation must be a priori. That is, the property must be one that the subject could (in principle) ascribe to the object solely on the basis of an understanding of the representational mode of presentation in question, and it should involve nothing more than is ascribable on that basis.

This condition on the relation between representational and nonrepresentational modes of presentation presupposes a commitment to what we might call thin properties--those with respect to which we know or can infer all there is to know about their intrinsic (nonrelational) nature merely in virtue of understanding the predicates that express them. Thin properties, then, confer no empirically discoverable essence or nature on the objects that instantiate them.⁸ Such properties, of course, are not necessarily phenomenal properties. Other types of thin properties include (among others) mathematical properties, (arguably) functional properties, such as being a telephone, (some) observable properties, such as being a medium sized object within reach, common sense properties, such as being a hill or an island, and common sense psychological properties, such as being angry or brave. Indeed we could say that thin properties correspond not to natural kinds but to "definable kinds," were we willing to take a sufficiently relaxed view of definition.

For a subject who believed that a fortnight was a period of ten days, for example, the representational mode of presentation (description) in question would not correspond a priori to the property of being a period of fourteen days. So it would not be in virtue of this property of a fortnight that the subject's term picked it out.⁹ Of course, the route from the linguistic expression to the referent may involve many representational modes of presentation, each with its corresponding property. If so, then reading down the path from top to bottom, each link must be either an a priori connection or an a posteriori connection grounded in the beliefs of the particular subject in question. And the reason is obvious. The route to the referent must explain how the subject's term succeeds in picking out the object in question.

Consider the analogy between the pain/C-fiber firing example and the Morning Star/Evening Star example. (See diagrams (1) and (2)). In both cases we have two linguistic descriptions that are coreferential but not coreferential a priori. In the case of 'the Morning Star' and 'the Evening Star' this is unproblematic. There are two properties corresponding to the two descriptions of Venus--the property of being the last heavenly body visible in the morning and the property of being the first heavenly body visible in the evening--properties that could have been instantiated by different objects. Thus the fact that one could fail to realize that a single object had both properties explains how one could believe--and be rational in believing-- something of Venus under the description corresponding to the first property ('the last heavenly body visible in the morning') and fail to believe it or believe the contrary under the description corresponding to the second ('the first heavenly body visible in the evening').

Now there is no difficulty about the modes of presentation (either representational or nonrepresentational) that tie 'my C-fiber firing' to its referent. The term 'C-fiber firing' is connected a priori with a certain causal role and the causal role (we can assume) is connected a posteriori (for the subject in question) with the perception of instrument readings and the like. The difficulty is with the modes of presentation of the token brain state as it is given "from the inside" to normal subjects, who determine that they are in pain without the aid of any special apparatus. There is a problem, however, regarding one of the nonrepresentational modes of presentation of the token state which is both the pain and the C-fiber firing. What is the property in virtue of which one has one's normal, first person access to such states? In diagram (2) as it stands, the route to the C-fiber firing goes through the property of being a

pain and the property of hurting or having a certain phenomenal feel. And these are, on the face of it, mentalistic properties. Can we, then, recast the diagram to provide a route that does not involve such mentalistic properties? It is evident that we cannot. No physical or neurophysiological property could play the role played by these ostensibly mentalistic properties in providing a route to the referent. For any such property, a perfectly rational subject can believe that he or she is in pain and not believe (or indeed disbelieve) what he or she would express in using a description connoting that property to characterize one of his or her internal states. We must then continue to appeal to a mental property--such as the property of being one's state that is hurtful at t --in order to produce an analogue of the Morning Star/Evening Star example.

But does the Morning Star/Evening Star example provide the general pattern for a posteriori identities, or is it merely one pattern among others, an option not a requirement? The suggestion that the pattern is optional seems clearly false. Recall that we want to rationalize the relevant beliefs, intentions, and actions of the subject who believes what he or she would express by saying "I am in pain" and "My C-fibers are not firing." Or, to put it another way, we want to say what the subject imagines or believes who is disposed to deny sincerely the identity of pain and C-fiber firing. And the answer must take the form of a specification of a possible world. First, suppose there were no such world (and nothing relevantly similar, such as a partial world or situation¹⁰) that was completely describable without contradiction and that captured what it was the subject imagined in imagining that pain was not identical with C-fiber firing. Suppose, that is, that every attempt to describe such a world ended in the revelation of a hidden contradiction. Then the identity would be a priori, contrary to our assumption. By the same token, a subject who failed to believe or disbelieved such an identity would be irrational. Thus we could not, for example, characterize the content of the relevant beliefs of the subject in terms of "impossible possible worlds" since such a characterization could not do justice to the subject's rationality, but would merely reveal the subject as irrational. What, though, would such a possible world be like, given that it could not be one at which pain = C-fiber firing since (by hypothesis) there are none? Such a world would have to be one at which (intuitively speaking) what the subject would mean in asserting the negation of the identity was actually the case. This means that there must be a world in which the properties that are definitive of the content (for the subject) of the two modes of presentation are not coinstantiated.

Even if this is how it works in the case of the Morning Star and the Evening Star, however, is it clear that it must work this way? The answer is yes. We need properties in order to characterize the way the subject takes the world to be, since we are specifying the content of the subject's belief, not merely something about its vehicle (such as its functional role). In other words, we are specifying a possible condition of the world, not merely a possible condition of the subject. We could derive the same conclusion by saying that we need properties of the world because, as we have already seen, local eliminativism is unacceptable and because it is the connection between properties and concepts that gives content to those concepts.

But if we need properties, is it clear that we need thin properties? Again the answer is yes. If the properties are to provide the content (for the subject) of the belief that he or she would express by saying "Pain is not identical to C-fiber firing," then their principle of individuation must be sufficiently fine-grained that they can be appealed to to explain the difference in cognitive significance between 'my pain' and 'my C-fiber firing'. Moreover, it must be finely grained enough to capture and explain the difference in cognitive significance of 'my pain' and any (explicitly) physicalistic expression, since for any such expression we can imagine a subject who believes what he or she would express using a sentence of the form 'My pain is F and it is not the case that P is F' where 'P' stands in for a referring expression couched in a purely physicalistic vocabulary. But this is just another way of saying that the property that provides the content for the subject of 'pain' is connected to the subject's use of the term a priori--i.e., the property's obtaining entails nothing more about the world than the subject could (in principle) infer on the basis of his or her understanding of the relevant representational mode of presentation. Thus we have the condition that the connection between the properties and the subject's expressions must be a priori in the relevant sense.

Is it really obvious, though, that there must be an a priori connection between the subject's representational modes of presentation and nonrepresentational modes of presentation (properties) in order to avoid local eliminativism? Consider the alternatives. First, the inferential roles (and their underlying functional realizations) of the expressions that figure in a subject's descriptions are not sufficient to provide their semantic content. Unless there is something to take us outside the circle of word-to-word connections, inferential roles, regardless of how extensive, could never provide more than

an uninterpreted calculus. The system of contents must contain some demonstrative element, and the question is where that element is to come from. As we have seen, bare causal connections to things in the world cannot, by themselves, provide the answer, since they are connections to which the subject need have no access.

But if inferential roles coupled with bare causal connections are not sufficient to provide content what is? A theory of meaning as use would be no help to the physicalist, since even if mental states are identified with physical states, the meanings of the physicalistic concepts will in turn be explained in terms of the contributions they make to our actions and practices--notions that themselves are mentalistic. Thus, if bare causal connections are unacceptable because they are connections to which the subject need have no access, then the explanation of content must involve a type of connection that rules out any such possibility. And this is exactly what justifies the claim that nonrepresentational modes of presentation (properties) must be individuated thinly and that the connection between such properties and the predicates that express them must be a priori in the sense defined.

We could not, then, fill out the analogy between the pain/C-fiber firing example and the Morning Star/Evening Star example simply by opting for an ostensibly mental property--e.g., being the state (or token event) of the relevant subject's that is hurtful at t --but claiming that it is identical with a neurophysiological property. Were we to do so we would not have provided an explanation as to how a rational subject could believe that he or she was in pain and in a state that was hurtful and not believe that he or she was in any relevant neurophysiological state. This is because there is no logically possible world at which the subject could be in a state that was hurtful and not be in the neurophysiological state in question.¹¹ Thus we could not describe the subject's beliefs, including the routes from the subject's referring expressions to their referents, in such a way as to satisfy (1)-(3).

We could satisfy (1)-(3), however, if we supposed that the first order property of being hurtful (which, we are assuming, is identical with a neurophysiological property) itself has two second order properties--a mentalistic property (e.g., the property of involving a certain phenomenal feel) in virtue of which the property of hurting has an a priori connection with the property connoted by 'my pain at t ' and a physical-functional property in virtue of which it could be picked out as the neurophysiological property that it is. Thus in order to satisfy (1)-(3), we are led to a dualism of second order properties. And if we

address this problem by identifying the ostensibly mentalistic second order property with another neurophysiological property, we have no way of avoiding an infinite regress.

We can summarize, then, by addressing the claim that Frege's constraint is satisfied by the distinction between the two concepts--the concept of a state that is hurtful and the concept of a neurophysiological state of type N. On this suggestion, no distinction is required at the level of properties: We simply have two distinct descriptions and concepts that pick out the same referent in virtue of the same (thick) property. First, as we have seen, we need to explain how the subject's expressions pick out the object in question. Now the property of being a neurophysiological state of type N explains how 'my neurophysiological state of type N' picks out the subject's C-fiber firing. But how does this property explain the fact that the subject's expression 'my state that is hurtful' picks out the same state? It must be in virtue of a different aspect of this property than the one in virtue of which the physicalistic expression picks out its referent. Thus the property of being a neurophysiological state of type N must itself have second order properties, and unless we postulate a mentalistic property at some higher order, the threatened regress will occur.

Second, in order to justify rationally the subject's beliefs, intentions, and actions, we have to be able to say how the world presents itself to that subject (condition (3c)). But to say how it presents itself to a subject is to specify a condition of the world; it is not a fact about the subject's concepts but about the contents of those concepts--i.e., about what they represent. Thus it will be a fact about the properties of the internal state in question. And these properties will have to be individuated finely enough to explain how the subject could be rational in believing what he or she would express by saying 'I am in a state that hurts and not in a neurophysiological state of type N'.

Third, we have to explain what it is in virtue of which the two concepts--the concept of being a state that is hurtful and the concept of being a neurophysiological state of type N--are distinct (at least as regards their cognitive significance). If it is claimed that they are distinct merely in virtue of a difference in functional roles and/or different causal chains to external objects, then we have local eliminativism. Since, however, there is nothing special about this local context, there is no principled way of avoiding eliminativism across the board, and this will seem to most too high a price for an objection to the property dualism argument.

Suppose it is admitted, however, that the difference in the cognitive significance of the corresponding descriptions 'my state that is hurtful' and 'my neurophysiological state of type N' for a subject is determined by the properties they connote. Then it is clear that the properties in virtue of which such descriptions pick out the objects they do must be given a priori. In other words, since the properties in virtue of which such descriptions pick out their referents provide the descriptions' (and the corresponding concepts') cognitive significance, those properties must be available to the subject solely on the basis of an understanding of the descriptions in question. Thus the properties connoted will be too thin to allow the kind of case proposed. That is, they will be too thin to allow a case in which the property of being hurtful is (unbeknownst to the subject) the property of being a neurophysiological state of type N and there is no genuinely mentalistic property not identical to a physicalistic property at the second or higher order.

III. BOYD'S ANTIKRIPKEAN ARGUMENT

It is sometimes claimed that the property dualism argument is vulnerable to an objection of Richard Boyd's to Saul Kripke's argument for dualism.¹² Kripke claims that we can imagine pain without any C-fiber firing (and vice versa) while acknowledging that we can, or think we can, also imagine water without H₂O even though they are the same. Kripke argues, however, that the cases of pain and water are not parallel. In the case of water and H₂O there is a plausible explanation of our thinking we can imagine water without H₂O when in fact we cannot (since if water is identical with H₂O, there is no such possibility). According to Kripke we confuse the case in which at some possible world water exists and H₂O does not (which we cannot imagine) with the case in which at some possible world a liquid other than H₂O plays the role that water plays at the actual world. On Kripke's account, however, there is no analogous explanation where pain and C-fiber firing are concerned. For there is no distinction between pain and the appearance of pain analogous to the distinction between water and the water role or the observable manifestations of water. As Kripke points out, whatever has the appearance of pain--i.e., whatever feels like pain--is pain. Thus, the intuition that we could have pain without C-fiber firing (and vice versa) stands as an objection to the identity theory.

Boyd objects that on Kripke's account, a possible explanation of our thinking (falsely) that we can imagine pain without C-fiber firing has been overlooked. Although there is indeed no distinction between pain and the appearance of pain, there is an obvious distinction between C-fiber firing and the way it manifests itself. Thus according to Boyd, we do have an explanation of our conviction that we can imagine pain without C-fiber firing: what we imagine is not pain without C-fiber firing but pain without the appearance of C-fiber firing. And this, Boyd maintains, is exactly analogous to the explanation of our believing that we can imagine water without H₂O. If the analogy holds, we could explain the apparent conceivability of pain without C-fiber firing, and hence the a posteriori character of their alleged identity, without postulating any mentalistic properties. Thus, according to Boyd, the fact that we seem capable of imagining pain without C-fiber firing is no more an objection to the identity theory than is the corresponding possibility regarding water and H₂O an objection to their identity.

Boyd's argument, however, does not supply the basis for skepticism about the original intuition (that we could have pain without C-fiber firing) in the way that Kripke's does for the intuition that we could have water without H₂O. Let us consider more carefully the nature of the alleged confusion on Kripke's account and its counterpart on Boyd's. On Kripke's account the confusion is over the question whether 'water' is a natural kind term or one that denotes anything that plays the water role. That is, it is over the question whether, with regard to any possible world, 'water' picks out the natural kind that plays the water role at the actual world, or, with regard to any possible world, it picks out whatever plays the water role at that world. However, once we fix the meaning of 'water' (as a natural kind term) and the fact that H₂O plays the water role at the actual world, the source of confusion disappears. Under these circumstances it is no longer even possible to think we can imagine water without H₂O. This is not so in the pain case. Suppose we make explicit to ourselves that we are using 'pain' not as a natural kind term, but in such a way that whatever feels like pain is pain. Imagine also that we make it perfectly clear to ourselves that 'C-fiber firing' is being used as a natural kind term (and not, say, as an operational concept such that anything that satisfies the standard tests for a C-fiber firing is one). Even with complete clarity on these points, we have no difficulty in imagining pain without C-fiber firing (and not merely in imagining pain without the standard evidence of C-fiber firing). If this is the case, however,

Boyd's claim--that confusion over what we are imagining contributes to the intuition that pain is not identical with C-fiber firing--does not undermine the Kripkean argument against the identity theory. For in this case, unlike the case of water, the intuition remains even when the source of the alleged confusion is eliminated. Since it does, the need to postulate different routes to the referent of 'pain' (in accordance with (1)-(3)) provides the premises on which the property dualism argument depends.

The general principle to which we appeal in this reply to Boyd's argument is that for a true identity statement, if neither of the referring expressions that flank the identity sign connote (pick out their referents in virtue of) contingent properties of those referents, the statement cannot be a posteriori. What we have shown, in effect, is that when we eliminate the contingency in the way water is picked out (by specifying that 'water' denotes a natural kind and that the natural kind that realizes the water role at the actual world is H₂O), we eliminate the a posteriori character of water's identity with H₂O. Since this is not the case for the identity statement connecting 'pain' and 'C-fiber firing', Boyd's argument fails. It is appropriate, then, that we turn to an objection to the property dualism argument that calls this principle (the so-called Semantic premise discussed below) into question.

IV. LOAR'S ARGUMENT AGAINST ANTIPHYSICALISM

Brian Loar's account of phenomenal states does not address the property dualism argument by name. It does, however, purport to address "an antiphysicalist line of reasoning that goes back to Leibniz and beyond" and of which Frank Jackson's knowledge argument and Kripke's antiphysicalist argument are instances (598). Moreover, (together with a similar argument of Block's) it provides the most sophisticated response available to this class of arguments and one that includes a detailed positive alternative to the antiphysicalist position.¹³ Let us, then, follow the usual practice in calling such antiphysicalist arguments conceivability arguments and consider whether Loar's criticism of the other members of this class are effective against the present argument for property dualism.¹⁴

Loar has both an interpretation of the assumptions behind conceivability arguments and a refutation of the arguments so interpreted. He has, in addition, a positive physicalistic

account of the meanings and referents of phenomenal terms such as 'pain' which, were it adequate, would undercut any conceivability argument for the postulation of mentalistic properties. I shall set out Loar's objections to conceivability arguments as he interprets them in this section and his positive views in section V. (It should be noted that whereas the property dualism argument concerns the identity of token events, Loar's discussion concerns the identity of properties. Since in both cases largely the same considerations apply, I shall refer to the difference explicitly only when necessary.)

Loar accepts Kripke's claim that the phenomenal concept of pain conceives of it directly and essentially but rejects the conclusion that pain cannot be identical with a physical property. According to Loar, the inference depends on an implicit assumption:

(Semantic premise) A statement of property identity that links conceptually independent concepts is true only if at least one concept picks out the property it refers to by connoting a contingent property of that property (600).

(Pairs of conceptually independent property concepts are just those that give rise to a posteriori property identity statements.) Loar's objection to the antiphysicalist position is that "a phenomenal concept can pick out a physical property directly or essentially, not via a contingent mode of presentation, and yet be conceptually independent of all physical-functional concepts" (600). Thus there could be a true statement of property identity linking a physical-functional concept and a conceptually independent phenomenal concept such that neither concept picks out the physical property in question by connoting a contingent property of that property. In other words, Loar's objection simply amounts to a denial of the Semantic premise.

Whether Loar's analysis of the antiphysicalist argument provides a fully adequate account of Kripke's or Jackson's position is a question I shall not try to answer. It would not, however, be an adequate characterization of the property dualism argument. Loar depicts the antiphysicalist commitment to the Semantic premise as based on the following intuition, which I shall call the intuition of transparency.

Phenomenal concepts and theoretical expressions of physical properties both conceive their references essentially. But if two concepts conceive a given property essentially, neither mediated by contingent modes of presentation, one ought to be able to see a priori—at least after optimal reflection—that they pick out the same property. Such concepts' connections cannot be a posteriori; that they pick out the same property would have to be transparent (600).

I shall defend a version of this intuition below. But one needn't accept this intuition in order to see the force of the Semantic premise. Rather, the Semantic premise emerges (in a weakened and modified form) as the conclusion of an argument.

As we have seen, the proponent of the property dualism argument begins with a commitment to the strong version of Frege's constraint (SVFC). This involves a commitment to rationalizing the subject's intentional states and actions—to providing them with a rational justification, in part by characterizing the world as the subject conceives it. This commitment is constitutive of the project of radical interpretation and intentional ascription. Furthermore, it involves a commitment to antielimativism. This is obvious where the global elimination of intentional states is concerned but less obvious where the eliminativism is local and part of an ostensibly antielimativist program.

The argument, then, to what I shall call the Weakened modified semantic premise is as follows. Assume that 'water = H₂O' is a true identity. And assume that it is knowable only a posteriori; in other words, the concepts 'water' and 'H₂O' are conceptually independent in Loar's sense. Thus there could be a perfectly rational subject who believed what he or she would express by saying 'Water fills the lakes and reservoirs' and also what he or she would express by saying 'H₂O does not fill the lakes and reservoirs'. There must, then, be a possible world—one describable in complete detail without a contradiction—which justifies this belief. The possible world that rationalizes and justifies the subject's beliefs, however, needn't be one at which water ≠ H₂O. After all, our commitment is to making the subject's beliefs rational. (By and large. We can ascribe irrationality, but only against the background of largely rational relations between intentional states.) There is no such commitment to showing that the beliefs are possibly true. If there are necessary truths known only a posteriori, then a rational subject could form

beliefs incompatible with them--beliefs that would be false at every possible world--and yet be rationally justified.

But what kind of world would justify the belief that the subject expresses at the actual world by saying 'water is not H₂O'? And what kind of world would justify a subject's (say Smith's) willingness to contribute to what is described as "research into the nature of water" which he believes necessary to sustain all life but not to contribute to what is described as "research into the nature of H₂O" which he believes is an extremely rare and inert substance, irrelevant to biological life? Clearly the notion of worlds at which Smith's beliefs are justified needs some explanation. There is a sense, after all, in which one's belief that a theorem of mathematics is false is justified by a world in which it is publicly shared by the most prominent mathematicians. The existence of worlds of this kind, however, does nothing to justify rationally one's belief, since its falsity is knowable a priori and hence is, in the relevant sense, irrational. Thus we cannot demonstrate Smith's rationality by finding worlds at which his beliefs are true, since there are none, and we cannot do so by finding worlds at which they merely seem true, since this is not sufficient for justification in the relevant sense. What then could we possibly hope to find?

It would be sufficient that there be a world (possibly with different physical laws) at which the terms 'water' and 'H₂O', though they are tied to the same (representational and nonrepresentational) modes of presentation as Smith's words, pick out two different substances with precisely the properties Smith ascribes to them. The answer, therefore, is that we are looking for possible worlds at which the narrow contents of Smith's beliefs are accurate or veridical, where narrow content is simply understood as the content that satisfies (all the relevant versions of) Frege's constraint--or, equivalently, the content that fully captures the cognitive significance of the subject's beliefs. (We say that the narrow contents are accurate or veridical rather than true because 'true' has long since been co-opted by proponents of broad content.) Indeed, this is all we could be looking for. The worlds that demonstrate the subject's rationality are the worlds that show that contents which fully reflect the cognitive significance of the subject's beliefs could all have been realized.¹⁵

Thus, Smith's belief need not be unjustified or irrational, even though there is no possible world at which water = H₂O. A world at which 'water' and 'H₂O' pick out different substances with the properties

that Smith ascribes to them will be sufficient. All we need to assume is that the properties that Smith associates with water--being colorless, odorless, tasteless, filling the lakes, etc., are instantiated by some substance other than H₂O.¹⁶ But this is just to say that Smith's term 'water' has its meaning and picks out its referent in virtue of

being associated with properties that are contingently connected with H₂O.¹⁷

If an identity is a posteriori, then, it might seem that there must be contingent modes of presentation (both representational and nonrepresentational) associated with at least one of the designating expressions that flank the identity sign. Suppose, however, that we consider Kripke's theory according to which the properties of having originated from a certain egg cell and having originated from a certain sperm cell are necessary properties of a person. Imagine an artificial fertilization process which includes a pair of egg cells named 'A' and 'B' and a pair of sperm cells named 'Y' and 'Z'. When the process is complete, one of the researchers involved says correctly that the person who originated from A (call her Eve) is the person who originated from Z. But this is an a posteriori identity linking conceptually independent concepts. (It is an empirical fact, discoverable only a posteriori, that A was combined with Z rather than Y.) Thus we have a violation of the Semantic premise (or its analogue for individuals) since the two descriptions in a true, a posteriori identity statement pick out Eve in virtue of their connoting two of her necessary properties. And this is evidently a straightforward counterexample to the Semantic premise, since even if Kripke is wrong about the necessity of origins for persons, we could define entities for which the thesis is correct.¹⁸ The way of handling the counterexample, however, is also straightforward. In place of Loar's version we should substitute

(Modified semantic premise) A statement of property identity that links conceptually independent concepts is true only if the concepts pick out the property they refer to by connoting contingently coextensive properties of that property.

Thus although there is no world at which Eve lacks either of the two properties connoted by the researcher's descriptions, there are worlds at which the properties are instantiated by other individuals than Eve. Thus the Modified semantic premise is satisfied and we have an appropriate explanation--one

that satisfies conditions (1)-(3)--of the a posteriori character of the identity. Thus we have the argument for (a slightly modified version of) Loar's Semantic premise.

Recall, however, that the claim was that there was an argument for a weakened version of the premise. In what sense is what we have derived weaker than the premise Loar rejects? Consider once again the identity 'pain = C-fiber firing' and assume that it is true. Assume also that 'pain' picks out its referent because pain has the property of being hurtful. Suppose now that the physicalist allows this point but claims that being hurtful is itself a neurophysiological property. Does this possibility violate Loar's version of the Semantic premise? The obvious answer is yes, since we have a statement of property identity that links conceptually independent concepts and is assumed to be true, while neither concept picks out the property it refers to by connoting a contingent property of that property. But is this case compatible with SVFC? Again the answer is yes. It is true that the situation as described so far does not explain how a rational subject could believe what he or she would express by saying "I am in pain" and what he or she would express by saying "I am in no relevant neurophysiological state." But the fact that the situation described does not yet explain the a posteriori character of the identity does not show that this aspect of the identity cannot be explained in a way that is compatible with everything that we have assumed. The a posteriori character would be explained if the second order property itself had two different kinds of third order properties--one in virtue of which it is the neurophysiological property it is and one in virtue of which it involves the feeling it does.

The point, then, can be put as follows. The strong version of Frege's constraint does not force a dualism of properties upon us at any particular level. It requires only that there be some level at which there is such a dualism of properties. Thus the semantic premise for which we have argued is not Loar's but

(Weakened modified semantic premise) A statement of property identity that links conceptually independent concepts is true only if the concepts pick out the property to which they refer by connoting contingently coextensive properties of that property, or contingently coextensive properties of a property of that property, etc.

V. LOAR'S ALTERNATIVE TO ANTIPHYSICALISM

Loar's positive position has two basic components. First Loar holds that we can have true identities like 'the property of being pain = the property of being a C-fiber firing', where the designating expressions flanking the identity sign connote conceptually independent concepts and where the identities are therefore a posteriori. And Loar believes that for any genuine psychological state, it is such an identity that captures the relation between that state and the subject's neurophysiological basis. Thus, for example, Loar rejects analytical functionalism. Second, Loar wants to maintain that this is possible even though neither of the designating expressions involved picks out the common referent by connoting a contingent property of that referent.

How does Loar reconcile these two claims? According to Loar, phenomenal concepts are type demonstratives--hence they pick out their referents directly. And Loar has an answer to the critic who says that

if the phenomenal concept is taken to discriminate some physical property, it does so via a phenomenal mode of presentation . . . the phenomenal concept does not pick out a physical state nakedly . . . But that conflicts with your assertion that phenomenal concepts refer directly, with no contingent mode of presentation (604).

His response is to say that phenomenal concepts have two kinds of noncontingent modes of presentation.

(1) A phenomenal concept has as its mode of presentation the very phenomenal quality that it picks out (604).

(2) Phenomenal concepts have "token modes of presentation" that are noncontingently tied to the phenomenal qualities to which those concepts point (604).

By (2) Loar apparently means that particular (token) feelings of pain can focus one's conception of the type of feeling to which those token feelings belong. Fundamentally, then, Loar's reply concerns phenomenal qualities or properties, and his claim is that the physicalistic can say exactly what the antiphysicalist would say--that the phenomenal property (which is a physical-functional property) picked out by a phenomenal concept is its own mode of presentation. As Loar says "the idea that one picks out the phenomenal quality of cramp feeling by way of a particular feeling of cramp . . . is hardly incompatible with holding that the phenomenal quality is a physical property" (604-605). And he adds that "the main point is by now more than obvious. Whatever the antiphysicalist has said about these cases the physicalistic may say as well" (604).

Loar's basic strategy is to point out that the antiphysicalist wants to say that there is no distinction between the phenomenal quality and its mode of presentation and to ask why the physicalistic should not say exactly the same thing. The point that this obscures, however, is that the physicalistic and the antiphysicalist have radically different reasons for making what is only superficially the same claim. The antiphysicalist makes this claim on the basis of what we might call the acquaintance sense of direct reference. The account is this. Visual sense-data (to take the clearest and most carefully worked out example), like other modes of presentation, are postulated to explain and describe the way the world presents itself to the subject. Thus visual sense-data have all and only the visual properties that seem to be instantiated. If Neo hallucinates a woman in red, then there is no physical object of his visual perception that explains the character of his visual experience. However, there is, on this view, a mental entity--a portion of his visual field, say--that actually has the shape and colors that seem to occur in the actual world.

Moreover, in line with the assumption of a strong analogy between one's visual field and such pictorial media as paintings and photographs, such pictorial properties are assumed to exhaust the properties of sense-data. Unlike a table, but like a picture of a table, the corresponding visual sense-datum has no hidden sides. Sense-data have all the properties they seem to have and they have only those properties; we cannot be mistaken in thinking a visual sense-datum has a property of the appropriately pictorial sort, and none of their properties go unnoticed. Thus we explain the way the world presents itself visually in experience by postulating a special class of mental objects that actually have

the visual properties that the world seems to instantiate. And we should notice explicitly the analogy between sense-data and the primarily descriptive Fregean modes of presentation with which we have been concerned. Both are intended to characterize the world as it presents itself to the subject in order to justify rationally the subject's beliefs, intentions, and actions.

These characteristics of sense-data make it clear why they were (and ordinary objects were not) appropriate referents of Russell's logically proper names. Russell appealed to definite descriptions to provide the representational modes of presentation necessary to solve the Frege problems that arise for the use of ordinary proper names in a range of contexts, including a posteriori identities. In such examples as 'Hesperus = Phosphorus' the difference in cognitive significance between 'Hesperus' and 'Phosphorus', the fact that one could be perfectly rational in believing what one would express in saying 'Hesperus is F and Phosphorus is not F', and the fact that the identity is not a posteriori are all explained by the fact that 'Hesperus' and 'Phosphorus' are associated with different descriptions (e.g., as above, 'the first heavenly body visible in the evening' and 'the last heavenly body visible in the morning' respectively). As Russell certainly recognized, however, this cannot be the whole story. If every referring expression got its connection to the world by being associated with a definite description, we would be caught in another infinite regress. And such a regress would clearly be intolerable; we would never get outside the circle of language-to-language connections to establish a connection between language and the world.

Such a regress is halted, on Russell's account, by the existence of logically proper names--that is, designating expressions whose only semantic function is to pick out their referents directly, without the mediation of descriptive content. And postulating sense-data as the referents of these logically proper names was a move ideally suited to bring this regress to a halt. First, since the sense-data in question are visual in nature, their connection to objects in the external world--sometimes alleged to involve a natural relation of resemblance--is radically different from that of linguistic expressions. Thus, because they introduce no further linguistic or descriptive content, they allow us to break the circle of language-to-language connections.¹⁹

Second, because like the images in paintings and photographs they have no hidden sides, they are their own modes of presentation. Therefore they stop another potentially infinite regress--this time of modes of presentation in general.

Third, given that logically proper names refer to sense-data, there are no a posteriori identities involving such names. Since there is no distinction between sense-data and their modes of presentation, they are not presented in virtue of any contingent properties that might have been instantiated by something else. Thus there is no possibility that two routes to the referent that in fact converge might have picked out different objects and hence no possibility of an a posteriori identity. And this is exactly what we find. If we think we refer directly to our own current sense-data, it seems obvious that we know whether we refer to two different images or to the same one twice. Moreover, this would be true of anything to which we could refer directly in the acquaintance sense. (And notice that this is precisely the sort of transparency that Loar disparages and that I claimed earlier that we would see emerge as the conclusion of an argument. Nothing I say, however, commits me to the existence of visual sense-data, as I shall make clear below.) The result is that logically proper names referring to sense-data could be used by Russell (in conjunction with his theory of descriptions) to solve Frege's problems while terminating what would otherwise be an infinite regress of descriptive contents. And Russell did so by appeal to a class of entities that raised no new Frege problems of their own.²⁰

Of course, as I have argued elsewhere, visual sense-data as understood by Russell do not exist.²¹ But what is crucial in the present context is Russell's strategy for reconciling direct reference with a solution to the relevant Frege problems: his limiting such direct reference to objects that are nothing over and above their modes of presentation. And there is nothing in my arguments against visual sense-data to prevent our treating pains and/or their phenomenal properties in Russell's way. That is, there is nothing to prevent our supposing that they, like visual sense-data according to Russell, are nothing over and above their modes of presentation. And this, of course, is exactly what we do normally suppose. We assume that at least in one sense of 'pain' or 'hurts', we only experience pain and it only hurts as long as we notice it and that if we believe that it hurts we cannot be mistaken.

The upshot is this. The antiphysicalist has a story to tell about how pains, understood as irreducible mental entities analogous to visual sense-data, could be nothing over and above their modes

of presentation. Thus the antiphysicist can explain direct reference to pains, so understood, without leaving any relevant Frege problems unsolved. Loar makes what are superficially the same claims: that reference to phenomenal properties is direct and that phenomenal properties are their own modes of presentation. And on this basis he claims that whatever the antiphysicist can say, the physicalistic can say as well. But the antiphysicist and physicalistic claims, though they are couched in the same language, are radically different. When Loar says that reference to pains or to phenomenal properties is direct and that they are their own modes of presentation, what he means is the following. The referring expressions that pick them out do so in a way that is unmediated by descriptive content. Thus the expressions do not connote properties of those phenomenal properties or pains. There will be a causal chain in virtue of which the linguistic expression is connected with its referent, but the connection is not mediated by representational or nonrepresentational modes of presentation distinct from the referent itself. There is in this account no counterpart of the antiphysicist's claim that pains or phenomenal properties are nothing over and above the way they are given to the subject from the first person or subjective point of view.

What, then, is the conclusion? The problem for Loar is that ordinary demonstratives used to pick out ordinary objects (and not, for example, sense-data) do raise Frege problems--even in cases where it is clear that the subject has no access to a descriptive expression that could replace the demonstrative in singling out the referent. In the two tubes problem, for example, David Austin imagines a subject who, capable of focusing his eyes independently, looks with each eye through a separate tube at a red screen before him. Since he cannot tell exactly how the two tubes are oriented, he wonders whether "that (referring to the red circular area that he is in fact seeing with his left eye) is identical with that" (referring to the circular area that he is in fact seeing with his right eye). These descriptions of the circular areas, however, are unavailable to the subject. This is because he cannot tell which area is seen with which eye--either because his ability to focus his eyes independently means that there is no unified visual field, or because he believes that he may suffer from a condition in which objects seen with the left eye appear on the right and vice versa.²² Thus even cases in which a demonstrative reference is irreplaceable with an identifying description raise Frege problems.

It follows that even cases of demonstrative reference to ordinary objects that are not via identifying descriptions are not direct in the sense in which reference to a sense-datum using a logically proper name would be. In contrast to the case of acquaintance, ordinary demonstrative reference, whether or not we call it direct reference, requires, as the existence of the Frege problems demonstrates, representational and corresponding nonrepresentational modes of presentation in order to satisfy SVFC and avoid local eliminativism. Thus Loar's talk of direct reference does nothing to show how we could do without representational and nonrepresentational modes of presentation in cases such as these, unless we were willing to pay the cost and embrace the eliminativist option.

Furthermore, the physicalistic cannot make sense of the idea that a neurophysiological state is its own mode of presentation. At best the physicalistic could say that some aspect of such a state provides its nonrepresentational mode of presentation, but in this it is like any other objective entity. As such it is available from any number of points of view, and there are an indefinite number of aspects of the state to which the subject of the pain has no access. Thus there is no counterpart of the assumption that pains have only the properties they seem to have or of the assumption that we cannot be wrong in ascribing them the phenomenal properties we do. As a result, the possibility of a posteriori identities arises (as it doesn't for sense-data) and with it the possibility of Frege problems generated by the fact that different routes can converge on the same referent in ways that can be established only by empirical investigation.

The conclusion, then, is that Loar's strategy fails. The physicalistic cannot say that our reference to pain or to phenomenal properties is direct in the same sense in which the antiphysicalist makes this claim. When the antiphysicalist says this, he or she means that our normal access to our own pains or phenomenal properties is via Russell's notion of acquaintance. And, as we have seen, this approach is compatible with SVFC. When the physicalistic says that our access is direct, this means merely that there is no descriptive mode of presentation of the state or property. But appealing to this fact to describe the mode of presentation of the physical property alleged to be identical with the characteristic feeling of pain involves the physicalistic in a dilemma. Assume that there is a representational mode of presentation available to the subject. (If not, if there is just a causal chain or process, then we have local eliminativism.) Then either the pain and the phenomenal properties are nothing over and above the mode of presentation and we have antiphysicalism or this is not the case and we lack a solution to the

Frege problems. And the suggestion that there might be some other mode of presentation (besides descriptive, causal, and via acquaintance) involves exactly the same dilemma. If the referents we pick out in virtue of such alternative modes of presentation are nothing over and above their modes of presentation, we have antiphysicalism, and if not, we have the Frege problems (and hence the property dualism argument) all over again. Thus I conclude that Loar's reply cannot be made to work.

VI. CONCLUSION

This account of the property dualism argument is an expansion of the version in "Curse of the Qualia." Though the current version adds substantially to the earlier one--particularly as regards the role of Frege's constraint, antieliminativism, and the threat to reductionists of an infinite regress, the conclusions only in this: While both versions of the argument yield the same disjunctive proposition, I have opted for a different disjunct in each case. The disjunction is that either there are irreducibly mentalistic properties or the connections between mentalistic and physical and/or functional concepts are conceptual and a priori. Since the second disjunct does not seem even remotely plausible for qualitative concepts and physical concepts, this comes down to the following: Either there are irreducibly mentalistic properties or analytic functionalism is true. In "Curse of the Qualia" I opted for the latter alternative. This now strikes me as less plausible than it did, and in this paper I have simply ignored analytic functionalism as a live option. Those who hold this view are entitled to point out that nothing I have said here provides a substantive reason for abandoning it.

If the property dualism argument is correct, then, (and the assumption about analytic functionalism is justified) we are committed to the existence of irreducibly mentalistic properties. (Nor would such properties supervene on the physical properties.) Does this mean that we must take such qualitative properties as being a pain, hurting, feeling like that, and so forth as beyond the reach of any sort of explanation or analysis? The answer is no. The property dualism argument requires mentalistic properties but not necessarily qualitative properties. Thus it is compatible with an attempt such as Michael Tye's to reduce the qualitative to the intentional.²³ And the implausibility of analytic functionalism does not automatically translate into an argument for the implausibility of such an intentionalistic reduction. This translation would only be available if the proponent of an intentionalistic reduction were also committed to an analytic reduction of the intentional to the physical or the functional. In the absence

of this further commitment, however, there is no reason why we could not treat such so-called qualitative states as pain as representational and attempt to illuminate their ostensibly qualitative character on the basis of their representational properties. (I have given a sketch of a nonreductive account of intentionality and consciousness elsewhere.)²⁴

The final issue is the relation of the property dualism argument to other recent conceivability arguments--Jackson's knowledge argument and Kripke's modal argument--and to Joe Levine's explanatory gap argument.²⁵ Though some are inclined to see these arguments as standing or falling on the basis of the same considerations, a number of reasons suggest that among conceivability arguments and those that appeal to similar assumptions, the property dualism argument has a special status. Jackson and his critics, for example, seem to have reached an impasse over the question whether a subject with a complete knowledge of all the physical facts but no experience of colors acquires a new piece of knowledge or merely a new set of skills when she first encounters a red object.²⁶ And although there are important objections to those critics of the knowledge argument who claim that only know how and not factual knowledge is acquired, these objections have not proved conclusive.²⁷ Moreover, in the light of the property dualism argument, we can see why this should be the case. The issue between Jackson and his critics is whether the chromatically deprived subject gains a new belief content in her first encounter with a red object. And it is plausible to think that the ultimate court of appeal on issues of this sort resides in the principles of radical interpretation that constitute and govern all content ascription. That is, the ultimate appeal is to appropriate versions of the principle of charity and of Frege's constraint, to the principles of theoretical and practical rationality, and so forth. And a typical instance of such an appeal would be the claim that the subject who opts to contribute to research described as "water research" and not to equally important projects described as "H₂O research" must have two modes of presentation of water--i.e., two distinct contents under which the same object figures in the subject's beliefs. By its very nature, however, the example that supports the knowledge argument seems to rule out an appeal to first principles where content is concerned. The general form of such an appeal is that content is required to rationalize what would otherwise be interpreted as irrational or self-defeating behavior. However, precisely because Jackson's subject has access to all the physical facts--including the facts about when and why normal subjects would use the color vocabulary and which such uses

would be correct--she will never be guilty of the kind of apparent practical irrationality that would ground an appeal to Frege's constraint or to the principle of charity. Thus Jackson's argument is constructed in such a way as to preclude the appeal to first principles that provides the most important ground for the ascription of content in the context of the property dualism argument.

As a candidate for being the most fundamental argument in this area, Levine's explanatory gap argument seems equally problematic. Levine himself makes a case that the explanatory gap argument has this fundamental status when he attempts to explain the conceivability of a creature's occupying any given physical or functional state and its lacking any sort of qualia in terms of the lack of an explanation of the nature of the qualia in physical and functional terms.

It is because [my italics] the qualitative character itself is left unexplained by the physicalistic or functionalist theory that it remains conceivable that a creature should occupy the relevant physical or functional state and yet not experience qualitative character.²⁸

But this surely gets things backwards. Even if we had an explanation of the qualitative character of pain, say, in physical or functional terms, it would remain conceivable that a creature should have the physical or functional states in question while lacking the qualitative experience and have the qualitative experience while lacking the physical or functional states in question--or indeed any relevant physical or functional states whatsoever. This is so for the same reason that it is conceivable that H₂O should fail to produce the macro-level properties of water on Earth and that those macro-level properties might exist on the basis of a different microstructure or none at all. There are, after all, possible worlds at which the laws of nature are different, and logical or conceptual possibility is what ultimately governs the relevant distinctions between what is conceivable and inconceivable, a priori and a posteriori. It is the lack of an analysis of qualia in terms that would make it a suitable explanandum of a causal/physical explanation that is crucial, and not the lack of an explanation itself. Thus in this domain it is conceivability that is basic, and this fact is amply reflected in the structure of the property dualism argument.

Finally, although Kripke's modal argument seems most closely related to the property dualism argument, even here the latter seems to provide some advantage. In its appeal to Frege's constraint and to antieliminativism, the property dualism argument supplies a grounding for what some of Kripke's critics (e.g., Loar) have seen as unmotivated and unsupported intuitions. By locating the basic issues in the theory of the ascription of content rather than in the logic and metaphysics of modality, the property dualism argument grounds the intuition of transparency and a version of the Semantic premise in a way that Kripke's discussions thus far have not. Thus in this domain it is the property dualism argument with which I believe physicalists will have to come to terms.

If the property dualism argument is understood as a conceivability argument, however, it is likely to be objected that conceivability is not always a reliable guide to possibility and that there has been no argument that it is so here. But this would be to misconceive the nature of the argument. The argument is that if pain were identical with C-fiber firing this identity would be a posteriori and one could be rationally justified in believing something that one would express in a statement of the form 'Pain is F and C-fiber firing is not F.' But making coherent sense of the possibility of such a rationally justified belief requires that we postulate either irreducibly mentalistic states or of irreducibly mentalistic properties of first or higher order. And we are committed to the existence of the entities required to make coherent sense of our other commitments. Thus we are committed to the existence of irreducibly mentalistic properties.²⁹

STEPHEN L. WHITE

Tufts University

Notes

1. On such a supposition, the event of Smith's being in pain at t --i.e., the event of his C-fibers firing at t --is what Donald Davidson has called an unrepeatable or dated particular. See "The Individuation of Events," "Events as Particulars," and "Eternal vs. Ephemeral Events," in his Essays on Actions and Events (Oxford: Clarendon Press, 1980), pp. 163-203. For the view of events as property exemplifications see, Jaegwon Kim, "Events as Property Exemplifications," in Myles Brand and Douglas Walton eds., Action Theory (Dordrecht/Boston: Reidel, 1976), pp. 159-177 and Alvin Goldman, A Theory of Human Action (Englewood Cliffs, NJ: Prentice Hall, 1970), Ch. 1. This alternative view of events is addressed in the discussion of property identities in sections IV-V.

2. Stephen Schiffer, "The Basis of Reference," Erkenntnis 13 (1978), 171-206, at 180. See also Brian Loar, Mind and Meaning (New York: Cambridge University Press, 1981), pp. 99-100 and Christopher Peacocke, Sense and Content (Oxford: Clarendon Press, 1983), p. 109.
3. On nonconceptual content see Christopher Peacocke, "Does Perception Have a Nonconceptual Content?" Journal of Philosophy 98 (2001), 239-264.
4. Schiffer's version of Frege's constraint suggests that modes of presentation belong on the side of language and content. Brian Loar sometimes means by 'modes of presentation' representations such as descriptions or concepts and sometimes the things connoted or expressed by such representations, such as properties. See Loar, "Phenomenal States," in Ned Block, Owen Flanagan, and Guven Guzeldere eds., The Nature of Consciousness (Cambridge: MIT Press/Bradford Books, 1997), pp. 597-616, at 600. (Page references in the text are to this paper.)
5. Donald Davidson, "Radical Interpretation" in his Inquiries into Truth and Interpretation (Oxford: Clarendon Press, 1984), pp. 125-139. David Lewis, "Radical Interpretation" and "Postscripts to 'Radical Interpretation'" in his Philosophical Papers, Vol.1 (New York: Oxford University Press, 1983), pp.108-121.
6. See my "Narrow Content and Narrow Interpretation," The Unity of the Self (Cambridge: MIT Press/Bradford Books, 1991), Ch. 2. Condition (3) and the subsequent explanatory remarks constitute, in effect, a commitment to a thoroughgoing internalism in one important sense of the term--a sense motivated by the kinds of intuitions that give rise to Frege's constraint and such notions as internal justification and the subject's point of view.
7. This argument was first presented (in a much more compressed form) in my "Curse of the Qualia," Synthese, 68 (1986), 333-368, reprinted as Chapter 3 of The Unity of the Self and in Block, Flanagan, and Guzeldere, pp. 695-715. It was inspired by an argument of J. J. C. Smart's which he attributes to Max Black, though a number of the important points in the present argument are not found in Smart's or, indeed, in my own earlier version. See Smart, "Sensations and Brain Processes," Philosophical Review, 68 (1959), 141-156. The property dualism argument is discussed in Christopher Hill, Sensations: A Defense of Type Materialism (New York: Cambridge University Press, 1991), pp. 98-101, Jeff McConnell, "In Defense of the Knowledge Argument," Philosophical Topics 22 (1994), 157-187, Christopher Hill, "Imaginability, Conceivability, Possibility and the Mind-Body Problem," Philosophical Studies 87 (1997), 61-85, Brian McLaughlin, "Review of The Unity of the Self," Journal of Philosophy 94 (1997), 638-644, Lenny Clapp, "Senses, Sensations, and Brain Processes: A Criticism of the Property Dualism Argument," Southwest Philosophy Review 14, 139-148, 59th Annual Meeting, Katalin Balog, "Conceivability, Possibility, and the Mind-Body Problem," Philosophical Review 108 (1999), 497-528, and John Perry, Knowledge, Possibility and Consciousness (Cambridge: MIT Press, 2001) For closely related arguments see Colin McGinn, "How Not to Solve the Mind-Body Problem," Terence Horgan and John Tienson, "Deconstructing New Wave Materialism," and Brian McLaughlin, "In Defense of New Wave Materialism: A Response to Horgan and Tienson," all in Carl Gillett and Barry Loewer eds., Physicalism and Its Discontents (New York: Cambridge University Press, 2001), pp. 284-330.
8. Although there is an obvious connection to a proposal broached by Alvin Goldman according to which "properties Q and Q' are identical just in those cases where they are expressible by synonymous expressions" [A Theory of Human Action (Englewood Cliffs, NJ: Prentice Hall, 1970), p. 12], there is no commitment on the present account to a notion of synonymy, a concept likely to yield controversial results in those cases where it is needed most. Indeed, there is no commitment to a notion of meaning (in the sense of a systematic semantics for all the relevant terms) shared by all competent speakers of a language. What is at issue is not the meanings of the relevant terms, but the modes of presentation under which individual speakers hold their beliefs about the referents of those terms--in short, the requirements imposed by Frege's constraint. In this respect the appeal to thin properties differs in its motivation from David Lewis's appeal to properties as classes of possibilities in order to satisfy the requirements of a systematic semantic theory of a language. See "New Work for a Theory of Universals," in his Papers in Metaphysics and Epistemology (New York: Cambridge University Press, 1999), pp. 8-55, at 16-17.
9. For this kind of example see Tyler Burge, "Belief and Synonymy," Journal of Philosophy 75 (1978), 119-138. For the subject who is mistaken about the meaning of 'fortnight' the route might go through the description 'the period referred to by 'fortnight' in this community'. Nothing, here however, turns on the account of parasitic reference.

10. See, for example, Jon Barwise and John Perry, Situations and Attitudes (Stanford: CSLI Publications, 1999), reprint with additions of Jon Barwise and John Perry, Situations and Attitudes (Cambridge: MIT Press/Bradford Books, 1983)
11. The relevant notion of possibility here is logical or conceptual possibility--describability without contradiction. The appeal to logical possibility, however, must be understood correctly. As Kripke has argued, possible worlds are ways the actual world could have been. Moreover, in moving from the actual world to its possible alternatives, we keep our language fixed. In the present context this means that if the property of being hurtful, say, is identical with some neurophysiological property, then there is no logically possible alternative to the actual world at which they are distinct. Thus the possible worlds with which we are dealing are genuinely possible and not merely epistemologically possible. They correspond, then, with what at least some have meant by "metaphysically possible worlds."
12. See Hill, Sensations, p.101. Saul Kripke, excerpts from Naming and Necessity, Lecture III (Cambridge: Harvard University Press, 1980), pp. 144-155, reprinted as "The Identity Thesis," in Block, Flanagan, and Guzeldere, pp. 445-450; Richard Boyd, "Materialism without Reductionism: What Physicalism Does Not Entail," in Ned Block ed., Readings in Philosophy of Psychology, Volume One (Cambridge: Harvard University Press, 1980), pp. 67-106.
13. I shall consider Block's variation on this argument elsewhere.
14. In the conclusion I shall distinguish two different forms that conceivability arguments might take.
15. See my account of notional content in "Narrow Content and Narrow Interpretation." This appeal to narrow content plays a role analogous to that of Chalmers' appeal to what he calls the "primary intention" of a proposition. See David Chalmers, The Conscious Mind: In Search of a Fundamental Theory (New York: Oxford University Press, 1996), pp. 57-65. Chalmers' account of primary intentions is itself similar to an earlier account of mine of narrow content in "Partial Character and the Language of Thought," Pacific Philosophical Quarterly 63 (1982) 347-365, reprinted as Chapter 1 of The Unity of the Self. The difference between the two accounts turns on whether narrow content should be represented in terms of a certain kind of two-dimensional matrix or the diagonal of that matrix. This difference needn't concern us here, since the more recent account of notional content seems preferable to either in taking Frege's constraint as constitutive of narrow content. (Though the nonequivalence of my two accounts is obvious, some readers have evidently been misled by my failure to make this point explicit. See McLaughlin, "Review of The Unity of the Self.") So long as an account of narrow content satisfies Frege's constraint, however, nothing in the present context depends on the details.
16. The question may arise why, in addition to the thin properties and the a priori connection between representational and nonrepresentational modes of presentation to which we appealed in the discussion of the Morning Star/Evening Star example, we need a notion of narrow content. The answer is that in the Morning Star/Evening Star case and its pain and C-fiber analogues it is assumed that the subject refers successfully to an actual object. In the general case in which we rationalize a subject's intentional states, however, we cannot make this assumption.
17. It may well be incompatible with the basic laws of physics that a substance other than water could have all of its observable or macro-level properties or that water could have failed to have them. But even if it is, the basic laws of physics are not themselves conceptually necessary. Thus in the relevant sense of 'possible' it is not merely the case that there is a possible world in which the substance with the macro properties of water is not H₂O. There is also a possible world in which H₂O does not have the macro-level properties of water at the actual world.
18. I am grateful to Ned Block for pointing out the significance of this example.
19. Bertrand Russell, "The Philosophy of Logical Atomism," in Robert C. Marsh ed., Logic and Knowledge (New York: Putnam, 1971), pp. 175-281, at 201.
20. These semantic arguments, which are completely independent of any prior epistemological commitments, were never (to my knowledge) presented explicitly by Russell. As R. M. Sainsbury's discussion indicates, however, there are good reasons to suppose that Russell was committed to all of the premises of the arguments advanced above. See Russell (London: Routledge and Kegan Paul, 1979), pp. 76-88, esp. 87-88.
21. "Consciousness and the Problem of Perspectival Grounding," presented at the Workshop on Consciousness Naturalized, Certosa de Pontignano, Siena, May 28, 1999 and "Subjectivity and the Agential Perspective," forthcoming in Mario de Caro and David Macarthur eds., Naturalism and Its Prospects, Harvard University Press.

22. In addition to Austin's discussion of the two tubes problem, see his discussion of Sarah, the pharmacist-astronaut, both of which are found in What's the Meaning of "This"? (Ithaca: Cornell University Press, 1990), pp. 20-25 and 42-51.
23. Michael Tye, "A Representational Theory of Pains and Their Phenomenal Character," in J. Tomberlin, ed., Philosophical Perspectives, Vol 9, (Atascadero: Ridgeview Publishing Co., 1990), pp. 223-239, reprinted in Block, Flanagan, and Guzeldere, pp. 329-340.
24. "Consciousness and the Problem of Perspectival Grounding."
25. Frank Jackson, "What Mary Didn't Know," Journal of Philosophy 83 (1986), 291-295, reprinted in Block, Flanagan, and Guzeldere, pp. 567-570; Saul Kripke, "The Identity Thesis;" and Joseph Levine, "On Leaving Out What It's Like," in M. Davies and G. Humphreys, eds., Consciousness (Oxford: Blackwell, 1993), pp. 137-149, reprinted in Block, Flanagan, and Guzeldere, pp. 543-555.
26. Jackson's critics include David Lewis, "What Experience Teaches," in J. Copley-Coltheart, ed., Proceedings of the Russellian Society, University of Sydney, 1988, reprinted in Block, Flanagan, and Guzeldere, pp. 579-595 and Laurence Nemirow, "Physicalism and the Cognitive Role of Acquaintance," in William Lycan ed., Mind and Cognition (Oxford: Basil Blackwell, 1989), pp. 490-499.
27. See Loar, "Phenomenal Concepts," pp. 607-608.
28. "On Leaving Out What It's Like," p. 548.
29. Earlier drafts of this paper were presented at the NYU Language and Mind Colloquium, April 4, 2000 and at the Workshop on Conceivability and Possibility, University of Fribourg, Switzerland, December 8, 2001. I am grateful to the audiences on both occasions. I have also benefitted from discussion of the property dualism argument with Jody Azzouni, Ned Block, Davor Bodrozic, Wei Cui, Martin Davies, Joseph Levine, Thomas Nagel, Martine Nida-Rumelin, Christopher Peacocke, Ullin Place, Mark Richard, Stephen Schiffer, Gianfranco Soldati, Richard Swinburne, Michael Tye, and Stephen Yablo.

