

## The Logical Status of Mind

GEORGE BEALER

**M**uch work in contemporary psychology and philosophy of mind proceeds reductionistically by likening, or by equating, the mind to something that is believed to be better understood: a system of behavioral dispositions, the nervous system, or a computational device.<sup>1</sup> To liken a puzzling kind of entity to something we better understand is often a wise explanatory strategy, and to identify one type of entity with another often yields considerable economy of theory. However, advocates of the prevalent reductionistic approaches to the mind have not at an earlier stage tried to do something that is always appropriate in theoretical inquiry. Namely, they have not early on tried to define what intuitively the subject at hand is supposed to be; they have not tried to give a definition of mind whose immediate purpose is simply to fit the clear-cut, intuitive examples of mental phenomena. This is the primary purpose of the present paper. When surveying these examples, one finds right on the surface purely logical features that are jointly necessary and sufficient for mentality. In this way, the clearcut, intuitive examples reveal that the distinction between the mental and the nonmental—unlike, say, the distinction between sugar and salt—is not naturalistic or empirical in character; rather, the distinction is so basic that it is logical in character. If correct, this is the sort of proposition that can be established before any empirical theorizing. And those who do not make use of this proposition lose the ability to demarcate clearly and precisely the very subject matter of psychology, namely, the mind.

With these bold claims before us, some conciliatory remarks are in order. When I suggest that it is appropriate early on in an inquiry to try to define basic notions of the subject at hand, I am not assuming that success is always possible. Many would hold that definitions of such things as heat, plant, or animal simply could not be given until thermodynamics, botany,

or zoology has reached a relatively advanced stage of development. However, at least provisional definitions of fairly basic notions are possible early on in some areas of natural science. For example, we have seen this rather often in mathematical physics: consider the classical Newtonian definition of acceleration as rate of change in velocity; this did not turn on particularly advanced developments elsewhere in the theory. And successful definitions of basic notions are commonplace at the outset of investigations in logic and mathematics. Only a philosophical dogma would exclude out of hand the corresponding definitional possibilities in psychological theory. The general point, then, is not that definitions are always possible early on; rather, the point is that, if such definitions are possible, surely they are desirable; therefore, it is always appropriate to try to find them. Psychological theory should be no exception.

The thesis that the distinction between the mental and the nonmental is logical in character is at odds with the philosophical naturalism so popular nowadays. An assumption of contemporary naturalism is that the natural sciences—physics, chemistry, biology—provide the immediate theoretical framework for psychological theory. But given the special affinity that reason has to logic, there is at least some plausibility in the thesis that it is logic that provides the immediate framework for characterizing certain basic mental notions. Again, it would be dogmatic simply to rule out this prospect.

The thesis that some interesting features of the mind might be discovered by purely a priori means is also out of step with the prevalent empiricism of our time. Yet the goal of psychology is not limited to the prediction and explanation of thought, experience, and behavior. Any acceptable psychological theory must also account for—or at least allow for—the fact that the psychologist has engaged in a psychological process: specifically, that the psychologist has arrived at his or her theory by rational means and, therefore, that it has epistemic merit.<sup>2</sup> To provide for this inevitable rational dimension, psychological theory must contain certain logically distinctive formal features. It should not be surprising that such formal features can be isolated a priori without the aid of empirical research. Reasonable empiricists should be able to admit that such transcendental loops might ensure a priori elements in their otherwise empirical theories. We shall see that such formal features are the ingredients needed for our purely logical analysis of mind.

Despite these conciliatory remarks, this sort of approach to the analysis of mind is unlikely to receive much support from the contemporary cognitive science community. In the present intellectual climate, the effort may be counted a success if others were simply to recognize in it a coherent alternative should their favorite approaches to the mind encounter difficulties.

## 1. INTENSIONAL LOGIC

When I say that the distinction between the mental and the nonmental is logical in character, I mean that the analysis of it can be stated entirely in terms belonging to logic; specifically, in terms belonging to intensional logic. In extensional logic, expressions having the same extension may be substituted for one another *salva veritate*. Intensional logic is that part of logic in which such a principle of substitutivity is not valid, at least *prima facie*. Consider a familiar example:

It is possible that some creature with a kidney is not a creature with a heart.

Every creature with a kidney is a creature with a heart, and conversely.

∴ It is possible that some creature with a kidney is not a creature with a kidney.

The premises are both true; the second premise ensures that 'creature with a kidney' and 'creature with a heart' have the same extension, and the conclusion arises from the first premise by substituting the first of these expressions for the second. Yet the conclusion is false. Hence, a *prima facie* counterexample to the substitutivity principle of extensional logic.

Working in the tradition of Gottlob Frege and Bertrand Russell, we may account for intensionality in logic by means of intensional entities. Intensional entities are the kind of extralinguistic entity that can be distinct from one another even when they are equivalent. Properties, relations, propositions, concepts, ideas, etc., are examples. In the above argument, the 'that'-clauses

that some creature with a kidney is not a creature with a heart

and

that some creature with a kidney is not a creature with a kidney

are singular terms denoting propositions. Because these propositions are both false, they have the same truth value and, hence, are equivalent. Nevertheless, they are distinct and can therefore have different properties. In particular, because the former proposition would have been true had biological evolution taken an appropriately different course, it has the property of being possible; because the latter proposition actually contradicts a law of logic, it does not have the property of being possible. Thus, the argument is invalid. Here, then, is a typical example of how intensionality in logic can be explained by means of intensional entities. Of course, various logicians and philosophers have tried to explain intensionality in logic without appealing to intensional entities. However, the known attempts to do so, including "syntactic" attempts like that of Carnap, have all run into one difficulty or

another. I submit that every attempt will upon critical examination be seen either to fail outright or to appeal covertly to intensional entities.<sup>3</sup>

There have been two prevalent conceptions of intensional entities in the history of logic and philosophy. On the first conception, intensional entities are taken to be identical if and only if they are *necessarily equivalent*. Thus, beyond the requirement of necessary equivalence, this conception just on its own imposes no further requirements on what it takes for a definition of an intensional entity to be correct. For example, each of the following candidate definitions taken from contemporary philosophy could be a correct definition as far as this conception of intensional entities is concerned:

- (a)  $x$  is grue if and only if  $x$  is green if examined before  $t$  and blue otherwise.
- (b)  $x$  is green if and only if  $x$  is grue if examined before  $t$  and bleen otherwise.<sup>4</sup>

The second conception of intensional entities imposes far stricter conditions on what it takes for a definition to be correct. On this conception, when an intensional entity is defined completely, the result is *unique* and *noncircular*. In this example, (a) is certainly a correct definition in view of its stipulative character. Therefore, on the assumption that complete definitions are unique, green must show up somewhere in the definition of grue either as a defined or as an undefined term. However, on the assumption that complete definitions are never circular, green cannot in turn be defined in terms of grue. Therefore, even though (b) provides us with a necessary equivalence, it cannot be a correct definition. Necessary equivalence is a necessary condition but not a sufficient condition for the identity of intensional entities.

The first conception underlies the currently popular possible-worlds treatment of intensional entities, and it also underlies Alonzo Church's "Alternative (2)" formulation of Frege's theory of senses.<sup>5</sup> This conception is particularly well suited to the treatment of the modalities (necessity, possibility, etc.), probability statements, counterfactuals, and so forth. The second conception has a rather livelier history. It underlies Leibniz's doctrine of simple and complex properties and Russell's doctrine of logical atomism. And when intensional entities are identified with ideas (concepts, thoughts), we see that this conception was adopted at least implicitly by nearly all modern philosophers from Descartes and Locke through Kant. For example, it is evident in the distinction made by Descartes and Locke between simple and complex ideas, and it underlies Kant's original definition of analyticity. Finally, this conception underlies Alonzo Church's "Alternative (0)" formulation of Frege's theory of senses.<sup>6</sup>

Some people who are friendly to the first conception of intensional entities might doubt the legitimacy of the second conception. However, just

as the first conception matches the intensionality present in the logic for the modalities, probability statements, and counterfactuals, the second conception matches the intensionality present in the logic for intentional matters—belief, desire, decision, assertion—and the intensionality present in the logic for logical truth and validity, provability, natural language semantics, and epistemic appraisal. Consider a sample argument involving logical truth:

It is logically true that all triangles are triangles.

It is necessarily true that all and only triangles are trilaterals.

∴ It is logically true that all triangles are trilaterals.

Given the first conception of intensional entities, the second premise of this argument would entail that being a triangle is identical to being a trilateral and, in turn, that the proposition that all triangles are triangles and the proposition that all triangles are trilaterals are identical. And if this is true, the conclusion of the argument would follow logically from the two premises. But intuitively the conclusion does not follow, for the proposition that all triangles are trilaterals is only a truth of geometry, not a truth of logic. The type of intensional entities that are the primary bearers of the property of logical truth are more finely distinguished than those provided by the first conception. The finely distinguished intensional entities provided by the second conception fill the bill. To see fine-grained intensions at work in natural language semantics, notice that

‘All triangles are trilaterals’ means in English that all triangles are trilaterals.

is true, whereas

‘All triangles are trilaterals’ means in English that all triangles are triangles.

is false. To see the call for fine-grained intensions in epistemic appraisal, notice that

The proposition that all triangles are triangles requires no epistemic justification.

is true, whereas

The proposition that all triangles are closed figures whose internal angles sum to 180° requires no epistemic justification.

is false. And for an example involving provability: it is provable in elementary logic that all triangles are triangles but not that all triangles are trilaterals.

Now there exist other conceptions providing intensional entities that, though more finely distinguished than those provided by the first conception,

UNIVERSITY MICROFILMS  
 SERIALS ACQUISITION  
 300 N ZEEB RD  
 ANN ARBOR MI 48106

are not as finely distinguished as those provided by the second conception. Virtually all of what I have to say below would hold if one of these intermediate conceptions were to serve the logical functions I have ascribed to the second conception. However, I submit that a consideration of the full range of examples inevitably forces one to this conception. Alonzo Church arrived at this conception by consideration of examples from intentional logic.<sup>7</sup> Nonintentional examples would also suffice for this purpose. Consider purely logical examples such as the following taken from the logic for the relation of logical validity (consequence):

- (1) Given that everything is self-identical, it follows validly by exactly one application of the principle of double negation that everything is not not self-identical.

It will not do to treat this example "syntactically":

- (2) Given the English sentence 'Everything is self-identical', the English sentence 'Everything is not not self-identical' follows validly by exactly one application of the principle of double negation.

because syntactic treatments fail the Langford-Church translation test.<sup>8</sup> Thus, translating (1) and (2) into a foreign language, say, German, we obtain two sentences that patently report quite different information:

- (1') Gegeben, dass alles selbstidentisch ist, folgt es stichhaltig durch genau eine Anwendung des Doppelnegationsprinzips, dass alles nicht nicht selbstidentisch ist.
- (2') Gegeben den englischen Satz 'Everything is self-identical', folgt der englische Satz 'Everything is not not self-identical' stichhaltig durch genau eine Anwendung des Doppelnegationsprinzips.

For example, whereas (1') reports something known by every German logic student, (2') reports something known only by people with knowledge both of some elementary logic and of the English language. Purely logical examples like (1) thus cannot be plausibly treated by means of syntactic entities. But their treatment requires extralinguistic "semantic" entities that nevertheless have a structure (or logical form) very much like that of syntactic entities. The fine-grained intensions provided by the second conception answer perfectly to this requirement.

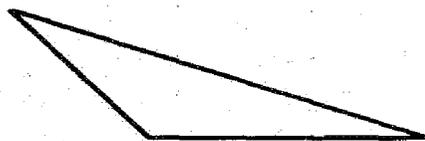
Intensional entities play an ultimate role in the objective, nonarbitrary categorization and identification of objects; in the objective description and theoretical explanation of change; and in the constitution of experience. But not just any intensional entities can play these important roles; the ones that can are said to be *qualities* and *connections*. Among the myriad intensional entities it is qualities and connections that determine the logical, causal, and phenomenal order in reality.<sup>9</sup> Now when qualities and connections are com-

bined by means of appropriate fundamental logical operations, sooner or later one comes to *conditions*. Conditions are the sorts of things that can be said to obtain or to be so.

Intensional entities that are neither qualities, connections, nor conditions are ones that pertain primarily not to the world but instead to thinking and to reason, taken broadly. Such intensional entities are called *concepts* and *thoughts*. Consider the example of green and grue mentioned earlier. Whereas green is a genuine quality (specifically, a sensible quality), grue is only a concept (i.e., the concept expressed in English by the complex expression 'green if examined before *t* and blue otherwise'). As such, grue plays no ultimate role in the objective, nonarbitrary categorization and identification of objects; nor does it play an ultimate role in the description and theoretical explanation of change; nor does it play an ultimate role in the constitution of experience. Nevertheless, like other concepts, grue can play a role in matters of thinking and of reason. Now, from a purely logical point of view, the difference between qualities, connections, and conditions on the one hand and concepts and thoughts on the other is that qualities, connections, and conditions conform to the first traditional conception; i.e., qualities, connections, and conditions are identical if and only if they are necessarily equivalent. However, though necessary equivalence is a necessary condition for the identity of concepts and thoughts, it is not a sufficient condition. For concepts and thoughts conform to the second traditional conception.<sup>10</sup>

Consider an example involving shape. Take the following figure:

( $\alpha$ )



What shape is ( $\alpha$ )? In answer to this question, one might say that ( $\alpha$ ) is a triangle. Or one could equally well say that ( $\alpha$ ) is a trilateral. Each of these answers suffices to inform us of its shape. The reason for this is that, intuitively, the quality of triangularity and the quality of trilaterality are the very same quality. They are how it is with ( $\alpha$ ) in regard to shape. Though the concept of being a triangle and the concept of being a trilateral are distinct, they correspond to the same quality of things in the world. The reason is that qualities, unlike concepts, are identical if there is no possibility of an object having one of them but not the other. Qualities (and connections) are what fix the actual conditions in the world, and they do not exhibit distinctions finer than necessary equivalents. Concepts, on the other hand, pertain to thinking and to reason; it is here that finer intensional distinctions show up. Thus, we arrive at the accompanying picture of the types of intensional entities.

UNIVERSITY MICROFILMS  
 SERIALS ACQUISITION  
 300 N ZEEB RD  
 ANN ARBOR MI 48106

	0-ary intensions	1-ary intensions	2-ary intensions	...
conception 2 (ideas)	thoughts	1-ary concepts	2-ary concepts	...
conception 1	conditions (states of affairs)	qualities	2-ary connections	...

Given this two-tier picture, it remains to account for how conception 2 intensions are related to conception 1 intensions. On realistic accounts, conception 2 intensions are constructed ultimately from qualities and connections (plus perhaps subjects of singular predication) by means of certain fundamental logical operations—singular predication, conjunction, negation, existential generalization, and so forth. So, for example, the thought that *x* spins is the outcome of predicating the quality of spinning of the object *x*. [That is, the thought that *x* spins = Pred(spinning, *x*).] A primary advantage of this account is that it includes a solution to the problem of representation: ideas succeed in representing things in the world because they actually are constructed from things in the world; nothing is needed in the account of how thoughts and concepts represent except the fundamental thought-building operations (like singular predication, conjunction, etc.) and objects in the world, including qualities and connections.<sup>11</sup> Notice that according to this theory, qualities and connections are not radically different from thoughts and concepts; indeed, they are just the limiting case. That is to say, they are those intensions that cannot be analyzed (defined) in terms of the fundamental thought-building operations plus other intensions. More precisely, they are in the range of none of these operations, and accordingly, they have no (complex) logical form.

But why should someone accept this full two-tier logical theory? Besides answering to a large family of logical and conceptual intuitions, the full theory is required for a long list of important theoretical tasks.<sup>12</sup> Suppose, however, we encounter someone (e.g., a nominalistically inclined, naturalistic empiricist) who resists this sort of logical theory on two counts: first, logical and conceptual intuitions do not qualify as legitimate data (evidence) because sense experience is the only legitimate source of data; and, second, in this person's own theoretical work there is no need to accomplish those special tasks that require this rich formulation of intensional logic. I believe that these two claims can be overturned by means of certain "transcendental" arguments. Although it is not feasible to present these arguments in detail here, I will sketch them in broad outline.

Concerning the first claim, what is it that makes our opponent's ex-

treme empiricist theory of data a rational one? Why is it not a mere expression of preference or a mere biohistorical episode that someone pursuing knowledge need not take seriously? Any response that falls back on the same empiricist theory of data only invites the same question. For example, suppose the response invokes a "total" theory that is generated by and that includes the empiricist theory of data. Why accept this total theory rather than some (perhaps simpler) alternate total theory that is generated by and that includes some alternate theory of data? Is the empiricist's choice more than an expression of preference or a biohistorical accident? One way for the empiricist to try to escape this circle would be to assert something like, "By definition, an item is a datum iff . . ." or "'data' is definitionally equivalent in English to ' . . . '." But how is any such definitional assertion to be defended? If our opponent again attempts an exclusively empirical justification (say, in the guise of an empirically defended behavioral or causal theory of meaning), the same sort of question may be asked once more; no significant progress will have been made out of the circle. Why is the thus-enlarged collection of (allegedly empirical) assertions more than a biohistorical episode having no special epistemic merit? Why not accept some (possibly simpler) collection of assertions that affirms some alternate definition of data? At this ultimate stage of epistemological dialectic, our opponent has no alternative left but to invoke as data logical and conceptual intuitions concerning the nature of data (or the meaning of 'data') and the application of the concept to actual and hypothetical cases. But if the empiricist must admit these logical and conceptual intuitions as data, it follows that the original empiricist theory of data cannot be consistently maintained. Further, if the empiricist must admit these intuitions as data, it would be arbitrary to exclude as data other logical and conceptual intuitions including, in particular, those that support our logical theory.

Next consider our opponent's second claim. (Here I will extend a line of argument developed by George Myro in "Aspects of Acceptability.") To justify their scientific and philosophical theories and to criticize those of their opponents, naturalistic empiricists must inevitably invoke a principle of epistemic appraisal or acceptability. The following, which is extrapolated from the writing of W. V. O. Quine, captures what is at the heart of these principles: a theory is acceptable if and only if it is (or belongs to) the simplest overall theory that explains the data.<sup>13</sup> A second basic principle of naturalistic empiricism is that, when taken together, the natural sciences (plus perhaps extensional mathematics) comprise the simplest overall theory that explains the data. It follows from these two principles that a theory is acceptable if and only if it is (or belongs to) the overall theory that consists of the natural sciences (plus perhaps extensional mathematics). Notice, however, that the expressions 'acceptable', 'simplest', 'explain', and 'data' do not belong to the primitive vocabulary of this overall theory. Let us suppose that these

UNIVERSITY OF MICHIGAN LIBRARY

expressions are not definable in terms of that primitive vocabulary. In this case, the two basic principles of naturalistic empiricism—and all the powerful conclusions that depend on them—would not belong to the overall theory and, therefore, would not be acceptable according to naturalistic empiricism. This self-defeating consequence can be avoided only if the key expressions ‘acceptable’, ‘simplicity’, ‘explanation’, and ‘data’ are, contrary to our supposition, definable within the overall theory. (We allow that some of these expressions and their definitions might have to be ramified à la type theory to avoid certain logical difficulties.) For this reason, naturalistic empiricism is forced to accept the thesis that these expressions are indeed so definable. But this thesis entails that the overall naturalistic theory must possess an apparatus for representing definitional relationships. Now this apparatus either would or would not be metalinguistic. If it were not metalinguistic, it would have to be (something like) one of the following:

By definition,  $Fx$  iff  $\dots x \dots$

It is definitionally true that  $Fx$  iff  $\dots x \dots$

The concept of being an  $x$  such that  $Fx$  = the concept of being an  $x$  such that  $\dots x \dots$

$F$ -ness = the property of being an  $x$  such that  $\dots x \dots$

(Mere material biconditionals would not suffice, for they lack the force of definitions.) But the logic for constructions of this type is clearly intensional, and the simplest explanation of intensional constructions is in terms of intensional entities. So on this option, naturalistic empiricists would be forced to invoke an ontology that, by their very own principles, is unacceptable. On the other hand, if the apparatus for representing definitional relationships were metalinguistic, it would have to involve a strong semantical term like ‘synonymous in  $L$ ’, ‘definitionally equivalent in  $L$ ’, ‘analytic in  $L$ ’, or some kindred term. But once again strict naturalistic empiricism would be thwarted. For generalized Quinean indeterminacy considerations show that strong semantical terms like ‘synonymous in  $L$ ’ cannot be defined using only terms belonging to the natural sciences and extensional mathematics. (See section 5 for more on this.) So on this metalinguistic option, naturalistic empiricists would again be committed to a definitional apparatus that, according to the central principles of their philosophy, is unacceptable. This is not to say that strong semantical notions like synonymy cannot be defined. They can be, but only in terms of intentional notions—along Gricean lines, for example. But considerations rather analogous to those leading to Quine’s indeterminacy thesis (see section 5 below) show that these intentional notions cannot be defined using only terms belonging to the natural sciences and extensional mathematics. At the core of these considerations is the fact that the logic for intentional matters is intensional (indeed, “hyperintensional”). However, as we have seen, the simplest explanation for intension-

ality in logic is in terms of intensional entities. So as in the case of a nonmetalinguistic apparatus, a metalinguistic apparatus for representing definitional relationships produces an ontological commitment to such entities.

Our conclusions so far are the following. Naturalistic empiricists are forced by their own basic principles to avail themselves of some apparatus for representing definitional relationships. But any such apparatus—metalinguistic or nonmetalinguistic—turns out to be unacceptable according to these very same principles. Thus, naturalistic empiricism in its strict form is a self-defeating philosophy and, hence, is unacceptable. Further, because any apparatus adequate for representing definitional relationships presupposes a logical theory that is ontologically committed to intensional entities, naturalistic empiricists have no choice but to weaken their basic principles to make room for this inevitable intensional ontology.

This line of argument indicates that a logical theory with an intensional ontology is inevitable, but can we go on to show that our full logical theory—with its ontology of qualities, connections, conditions, concepts, and thoughts—is inevitable? Do naturalistic empiricists have specific theoretical tasks to perform that force them to accept the full intensional logic? To show that they do, we must look to the specific notions they employ in the basic principles of their philosophy—in particular, the notions of simplicity and explanation. (At the close of the paper I will indicate how the full two-tier intensional ontology is also needed in an account of the notion of data.) Consider simplicity first. The only way to assess the simplicity of a theory is to express the theory using just primitive constants for qualities and connections; failing this, the simplicity of a superficial syntax can mask the true complexity of a theory.<sup>14</sup> Thus, in the defense of their second basic principle (stated earlier), naturalistic empiricists must invoke the ontology of qualities and connections. Next, consider explanation. Naturalistic empiricists hope to define this notion by means of a hypothetico-deductive account, which employs notions of logical validity and causal law. But how are these notions to be defined? Consider logical validity. A language-neutral, extralinguistic notion is what is needed for a *general* formulation of the hypothetico-deductive account of explanation. Such a notion may be defined along the following lines: a thought is logically valid iff<sub>df</sub> every thought having the same logical form corresponds to a condition that is necessary. And the notions of logical form, correspondence, and necessity are definable in the full intensional logic.<sup>15</sup> Finally, consider the notion of causal law. A standard regularity account is inadequate because it is unable to accommodate such phenomena as the antisymmetry of the explanation relation (*vis-à-vis* causal laws that are biconditional in form). Either the notion of causal connection (causal connections being inherently antisymmetric) or, at least, the general notions of quality and connection are required to accommodate such phenomena.<sup>16</sup> Once again, these notions are definable in the full

intensional logic.<sup>17</sup> In the above ways, then, the full intensional logic proves useful for dealing with the naturalistic empiricist notions of simplicity and explanation. At the same time, no essentially weaker theoretical framework appears to be adequate for all these purposes. In view of this, naturalistic empiricists appear to have no alternative but to modify the basic principles of their philosophy to make room for the full intensional logic.

Although these remarks have been highly schematic, they should at least make it credible that anyone inclined to espouse naturalistic empiricism—or, more generally, anyone with aspirations to theories comprehensive enough to account for their own acceptability or epistemic virtue—is forced eventually to incorporate the sort of intensional logic I have been advocating. It is this sort of logical theory that provides a framework rich enough for analyzing intentionality and mind.

## 2. INTENTIONALITY

An intentional phenomenon, according to Franz Brentano, is one that makes reference to, is directed upon, or is about other objects, perhaps even objects that do not exist. Brentano used this concept of intentionality to formulate his two-part thesis of intentionality:

- (1) All and only mental phenomena are intentional.
- (2) No purely physical phenomenon is intentional.

I will return to Brentano's thesis later. My goal now is to analyze intentionality—this special phenomenon of *aboutness*—without resorting to linguistic or spatial metaphors as Brentano did.

In the history of philosophy, the prevailing theories of intentional phenomena have been nonpropositional/nonrelational. On Brentano's theory, for example, when one judges that there exists a man, one does not stand in a relation to the proposition that there exists a man. Rather, one affirms or accepts men. Evidence of such nonpropositional/nonrelational theories are found in works ranging from Plato's *Sophist* and *Theatetus* to Russell's introduction for the first edition of *Principia Mathematica* and to his *The Problems of Philosophy*. These theories all suffer from a central shortcoming that is *logical* in nature: namely, they are unable to handle *general* statements concerning intentional phenomena.<sup>18</sup> For example, they are unable to account systematically and without circularity for intuitively valid inferences such as the following:

Whatever  $x$  judges  $y$  judges.

Whatever  $y$  judges  $z$  judges.

∴ Whatever  $x$  judges  $z$  judges.

Whatever  $x$  judges is true.  
 $x$  judges that  $A$ .

∴ It is true that  $A$ .

However, if we treat 'judges' as a two-place predicate and 'that  $A$ ' as a singular term, then such arguments are easily represented within the framework of quantifier logic:

$(\forall u)(xJu \rightarrow yJu)$

$(\forall u)(yJu \rightarrow zJu)$

-----  
 $(\forall u)(xJu \rightarrow zJu)$

$(\forall u)(xJu \rightarrow Tu)$

$xJ[A]$

-----  
 $T[A]$

However, in adopting this treatment we are led inevitably to the conclusion that 'judges' expresses a binary relation and that the range of this relation is made up of the sort of thing characteristically denoted by 'that'-clauses, namely, propositions.<sup>19</sup> And so it is that we arrive at the relational/propositional theory of judgment.

Nowadays, the adverbial treatment is perhaps the most popular non-relational treatment of judgment. On this treatment, 'x judges that  $A$ ' is rendered ' $(A(J))x$ '. Here ' $(A(J))$ ' is a compound monadic predicate, and the formula  $A$  functions as an adverbial phrase that modifies the monadic predicate  $J$  by fixing a relevant "way of judging" (intuitively, the way one judges when one judges that  $A$ ). On this treatment, a general sentence such as 'Whatever  $x$  judges  $y$  judges' must then be rendered ' $(\forall \phi)((\phi(J))x \rightarrow (\phi(J))y)$ '. Here the variable  $\phi$  takes formulas as substituends, and it functions as an adverbial phrase whose semantical values range over "ways of judging." Notice, however, that the construction ' $(\phi(J))x$ ' requires the adverbial theorist to admit that singular-judgment statements are true only if there is a relation—namely,  $[(\phi(J))x]_{x\phi}$ —that holds between individuals and "ways of judging." Further, since ' $(\phi(F))$ ' and ' $(A(F))$ ' are unfamiliar idioms, we are owed an account of their semantics. This can be done but only if the adverbial theorist in effect reintroduces the relational/propositional treatment in the metatheory. (For example, adverbial occurrences of  $\phi$  may be treated as fixing partial functions that take a certain kind of relational property to another kind of relational property: if  $F = [(\exists q)xR^2q]_x$  for some relation  $R^2$  and if  $p =$  the proposition expressed by  $\phi$  when it functions as a formula, then the partial function fixed by an adverbial occurrence of  $\phi$  takes  $F$  to  $[xR^2p]_x$ . Accordingly,  $[(A(J))x] = [xJ[A]]$ .) Consequently, nothing is gained by the adverbial approach, and hence, the essentially simpler relational/propositional approach is superior.<sup>20</sup>

Now each expression in the family 'thought', 'belief', 'judgment', and so forth, has at least three related uses. Each can be used to mean (1) a kind of intentional act, (2) the propositional object of the intentional act, or (3) a relation holding between persons performing the intentional act and the propositional object of the act. The nonrelation/nonpropositional theory acknowledges only the first of these three uses, the one for intentional acts. This forces the theory to give its account of intentionality in the inevitably opaque terms of intentional acts, making metaphor and circularity unavoidable. By contrast, the relational/propositional theory acknowledges all three uses and thus is free to analyze intentional acts in terms of the associated relations and their propositional objects. The following is an illustration of how easy these analyses can be:  $x$  performs the intentional act of thinking that  $A$  if and only if  $x$  stands in the relation of thinking to the thought that  $A$ .

However, I have said nothing yet concerning the intentionality of intentional acts, their directedness or aboutness. How does that arise? The answer is that it arises from the intensional entities, (e.g., from the thoughts) to which the person stands in the relation thought, belief, judgment, and so forth. After all, thoughts in the propositional sense are themselves things that are characteristically said to be about other objects; indeed, they are often said to be about objects that do not exist. The same thing holds not just for thoughts but for all ideas, concepts as well as thoughts. On this account, an intentional act can be said to be about other things for one reason only; namely, the intentional act consists in standing in a certain relation to an idea—a thought or concept—that can be said to be about things. (Intentional acts of the kind described with objectival constructions might seem to be an exception; section 3.3 shows why this is not so even if these constructions are taken at face value.) This is to say, an intentional act can be said to be about other things only secondarily through the idea that is the immediate object of the act. Ideas—the type of intensional entities provided by the second traditional conception—are the objects that can in a primary way be said to be about other things. (Qualities, connections, and conditions are not like this. They simply qualify, connect, or obtain; they are not *about* anything.)

With the intensional logic described earlier, we are able to catalogue the ways in which ideas—thoughts and concepts—can be about other objects, even nonexistent ones. And we can do this entirely in terms of fundamental logical operations on intensional entities: namely, singular predication, conjunction, negation, existential generalization, etc. For example, the most direct way in which a thought can be about an object occurs when the thought is the outcome of a singular predication and the object is its subject (i.e., the thought that  $Fx = \text{Pred}(F, x)$ , where  $x$  is the object,  $F$  is a 1-ary intension, and  $\text{Pred}$  is the logical operation of singular predication).

So far, then, we have the following conclusions. First, there are independent logical grounds supporting the relational/propositional theory of thinking. Second, using the relational/propositional theory, we are able to analyze intentional acts in terms of the associated relations and propositional or conceptual objects. Third, the intentionality (i.e., the aboutness) of an intentional act can be accounted for by the fact that the intentional act consists in standing in an appropriate relation to an idea—either a thought or a concept—that, given the right context, can be said to be about other objects, even objects that do not exist. Fourth, using the logic for thoughts and concepts, we can identify all the formal features that are at work in determining what in a given context an idea can be said to be about.

Yet the story is not complete, for there is an unsolved problem. Standing in just any relation to a thought or concept does not constitute an intentional act. Only certain very distinctive relations will do—relations such as thinking, believing, judging, remembering, perceiving, desiring, deciding, intending, etc. These relations, naturally enough, are called intentional relations. The problem is to give a noncircular definition of what an intentional relation is. If this problem can be solved, then the analysis of intentionality will be complete.

What makes this problem seem difficult initially is that there are many ad hoc relations (i.e., grue-like relations, which are not genuine connections) whose logical behavior is very much like that of genuine intentional connections. (Recall that connections are those special relations that, together with genuine qualities, fix the logical, causal, and phenomenal order of the world.) The key to the problem is to proceed in two stages. The first stage is to define what an intentional connection is. In the second stage, we then state what it takes for an ad hoc relation to be intentional; it is one whose analysis depends in a logically essential way on intentional connections. Naturally, there are many forms this kind of dependence can take, so there are many ways in which an ad hoc intentional relation can be said to be intentional. But this is a technical point not important to the central concerns in philosophy of mind. The core of our problem comes down to the problem of discovering what is logically distinctive about intentional connections.<sup>21</sup>

Given what I have already said, we know that intentional connections can connect individuals to ideas—thoughts or concepts. However, there are certain nonintentional logical connections that are like this too. Consider an example that arose earlier, namely, the operation of singular predication. This basic logical relation connects individual objects  $x$  and 1-ary intensions  $F$  to thoughts that  $Fx$ . Notice, however, that whenever this relation holds among three such items, it does so necessarily, not contingently. Thus, when this relation holds, we do not have what we may call a real phenomenon; we have instead a logically necessary condition. Intentional connections, by contrast, are typically not like this: when an intentional connection holds,

typically it does so contingently, not necessarily. This is why intentional connections give rise to genuine phenomena. Thus, in our definition we should require of intentional connections that they be able contingently to connect individuals to ideas—concepts or thoughts. I call this the *contingency requirement*.

For a second example of a nonintentional relation that can connect individuals to ideas, consider the relation of falling under a concept. This relation meets the contingency requirement: it can contingently relate an individual to a concept. (For example, it is contingent that Socrates falls under the concept of being more than five feet tall.) But notice that, whenever this relation holds between an individual and a concept, it must also hold between the individual and all necessarily equivalent concepts. Intentional connections, by contrast, are not like this. An intentional connection can contingently connect an individual to a concept independently of whether it connects the individual to necessarily equivalent concepts. Analogously, an intentional connection can contingently connect an individual to a thought independently of whether it connects the individual to necessarily equivalent thoughts. I call this feature *hyperintensionality*.<sup>22</sup>

There are a number of nonintentional, naturalistic relations that can contingently relate individuals to ideas. None of them, however, is a connection with the special additional feature of hyperintensionality. Consider an example.<sup>23</sup> Suppose that we come across what appears to be a rabbit's footprint in the snow. We might say of this little depression in the otherwise smooth surface that it makes it probable that a rabbit was here. But notice that, necessarily, if a particular  $x$  makes a thought  $y$  probable and  $y$  is necessarily equivalent to another thought  $z$ , then  $x$  also makes  $z$  probable. (That is, for each argument  $x$ ,  $x$ 's image under the rendering probable relation is closed under necessary equivalence.) However, nothing analogous is true of intentional connections. To see this, consider an example taken from Quine.<sup>24</sup> The thought that a rabbit was here is necessarily equivalent to the thought that an undetached rabbit part was here. Thus, necessarily, if a depression in the snow renders it probable that a rabbit was here, then the depression also renders it probable that an undetached rabbit part was here. However, a person can think (remember, desire, etc.) that a rabbit was here and yet fail to think (remember, desire, etc.) that an undetached rabbit part was here. Indeed, the person might at the moment be quite unable to grasp the latter thought. Now the broader philosophical point is this. The nonintentional, naturalistic relation of rendering probable does not have the capacity to cut those fine-grained intensional distinctions characteristic of conception 2: standing in this relation to one thought requires standing indiscriminately in it to all necessarily equivalent thoughts. Thus, in effect, this relation relates individuals not to thoughts singly but instead to equivalence classes of necessarily equivalent thoughts. Now there is a one-one map

that takes each such equivalence class to the condition to which the thoughts in the class correspond. Thus, in effect, this relation of rendering probable does little more than relate individuals to conditions. Intentional relations, by contrast, make discriminations finer than necessary equivalence, opening up the possibility of beings whose states are indistinguishable from the point of view of naturalistic (i.e., probabilistic) information flow but who nevertheless are in distinct states. (Of course, there might be hyperintensional naturalistic relations that are not real connections; some candidate examples are discussed in section 3.4.)

There are traditional philosophical views that intentional connections need not display both contingency and hypertensionality with respect to *every* individual and *every* idea. For example, on certain views God necessarily thinks all and only true thoughts, and certain inanimate objects (e.g., stones) necessarily think no thoughts. Further, some philosophers hold that anyone who thinks at all must think certain thoughts; for example, if I think at all, perhaps I must think I exist (at least as a transcendental unity of apperception). And other philosophers maintain that there are logically degenerate thoughts such that, if we think one of them, we must also think at least some other thought that is necessarily equivalent to it; for example, if we think that  $A$  and  $A$ , then we must also think that  $A$ . Finally, on some views, there are some thoughts that are impossible for any one person to think. For example, suppose that  $p$  is the first-person thought someone thinks when asserting 'I am thinking', that  $q$  is the first-person thought someone else thinks when asserting 'I am thinking', and that  $r$  is the result of conjoining  $p$  and  $q$ . Then, perhaps the privacy of a person's first-person thoughts makes it impossible for anyone to think  $r$  directly (i.e., without the aid of any intervening descriptive concepts). Now many people will wish to question some of these views. (For example, some will doubt any view implying that God's thoughts are necessarily determined and, hence, that He is unfree, and others will doubt the limitation on what is logically possible implicit in the metaphysical essentialism concerning stones.) In any event, even if each of these views is correct, it is clear that intentional connections can be contingent and hyperintensional for *some* individuals and *some* ideas.

The resulting logical picture is this. The family of genuine connections that can hold between an individual and some other item is quite distinctive to begin with. And the connections that can hold between an individual and an idea are particularly distinctive. Of the latter connections, all and only those that are intentional can be both contingent and hyperintensional.

Assembling the foregoing ideas, we arrive at the following analysis:

A connection is intentional if and only if it can contingently connect an individual to a thought or a concept independently of whether it connects the individual to a necessarily equivalent thought or concept.<sup>25</sup>

In symbols, a connection  $c$  is intentional iff<sub>df</sub>

$$\Diamond(\exists xyz)(\text{Ind}(x) \ \& \ \text{Idea}(y) \ \& \ y \approx_N z \ \& \ \Diamond(x, y \Delta c \ \& \ x, z \not\Delta c) \ \& \ \Diamond x, y \not\Delta c).$$

What is most distinctive about this analysis is that it is given entirely in terms of logic, specifically, the intensional logic I described earlier. (The ultimate primitive terms of the analysis are just those for intensional abstraction and the copula  $\Delta$ .) The view that emerges is that the intentional/nonintentional distinction is so basic that it is neither naturalistic nor empirical in character; rather, it is purely logical. The claim here is that all and only intentional connections have the indicated logical character; what is distinctive about us intentional beings is that we can stand in contingent connections to ideas independently of the necessarily equivalent ideas to which we might be connected. Within the tide of naturalistic information, we intentional beings exercise our capacity to be connected to subtly distinct aspects of that brute flow and, indeed, pursue our lives in these terms. It is thus that the phenomenon of intentionality emerges into the less subtle world of nature.

### 3. POTENTIAL COUNTEREXAMPLES

There are a number of potential counterexamples to this analysis of the concept of intentional connection. Although many appear promising at first, they can all be dealt with. Still, some of them are of philosophical interest in their own right. For this reason, as well as to impart a better feel for the analysis, I will explain why the best of these counterexamples fail.

#### (1) *Ad hoc relations*

There are countless ad hoc nonintentional relations that can be defined by tricks of elementary logic so that they have the properties of contingency and hyperintensionality. For example, let  $Rxy$  iff<sub>df</sub>  $x$  is more than five feet tall and  $y =$  the proposition that there are nine planets. Then the relation expressed by ' $Rxy$ ' holds between anyone more than five feet tall and the proposition that there are nine planets, and it does so contingently and without holding between such people and any necessarily equivalent propositions. Obviously this relation is not intentional, but just as obviously it is ad hoc and not a genuine connection. So it is not a counterexample to the proposed definition of intentional connection.

Naturally, there also are ad hoc intentional relations that lack contingency, hyperintensionality, or both. However, the second stage of our two-stage strategy takes care of intentional relations of this sort. In each case, the definition of the relation involves intentional connections in some logically essential way. For example, the relation holding between  $x$  and  $y$  such that  $x$  believes something that is necessarily equivalent to  $y$  does not have hyper-

intensionality. However, the definition of this nonbasic relation involves an intentional connection—namely, belief—in a logically essential way. So it qualifies as intentional at the second stage of the proposed analysis.

For a somewhat more interesting family of nonbasic relations, let us consider physical measure functions. These functions are nonintentional relations that contingently correlate individuals with numbers. For example, the relation holding between a parent and the number of his children, the relation holding between a physical object and the number of grams in it (i.e., the number of disjoint one-gram parcels in an exhaustive decomposition of it), the relation holding between a physical object and the number of ounces in it, and so forth. For these relations to be even *prima facie* counterexamples to our analysis, numbers would have to be identified with concepts. In this case, we should ask to what sort of things these number concepts apply. A broadly Fregean answer is surely the best: numbers are not concepts of *single* individuals (e.g., this parent or that piece of gold); rather, they are concepts of the abstract items (sets, concepts) that we employ for the purpose of thinking about individuals *collectively*. When we say that three is the number of  $x$ 's children, we are not suggesting that each (or any) of  $x$ 's children is three or three-ish. Rather, the set (or concept) of  $x$ 's children is three or three-ish; the set (or concept) falls under the numerical concept of being a set (concept) with three individuals in it. Now the details of this kind of approach are not important here; and, as is well known, there are many alternative analyses that differ from Frege's own in one way or another. What is important is that on any broadly Fregean analysis, physical measure functions turn out to be nonbasic relations defined exclusively in terms of fundamental formal relations (such as the logical relation of falling under a concept) and physical relations that hold exclusively among individuals (such as the parent relation or the relation of weighing the same).<sup>26</sup> Therefore, even though physical measure functions might have a superficial resemblance to genuine intentional connections (by virtue of the fact that they can contingently correlate individuals and intensions), they are significantly different. Upon closer examination they are revealed to be nonbasic relations definable in terms of underlying physical and formal relations that are not even in *prima facie* conflict with the proposed analysis of intentional connections.<sup>27</sup>

For a final example of nonbasic relations that bear a superficial resemblance to genuine intentional connections, consider *utterance-token meaning*—as in the relation holding between utterance tokens and what they mean in a community. This is a contingent hyperintensional relation: it can contingently relate a particular (namely, an utterance token) to an idea (namely, the meaning of the utterance token in the community) without relating the particular to any necessarily equivalent ideas. Is utterance-token meaning truly a connection? It hardly seems so. An utterance token and the relevant idea are not related to one another just on their own; the active intervention

of a third element is required, namely, the intentional activity of the thinking beings in the community. Not unless these beings make utterances with appropriate intentions and beliefs do utterance tokens become related to the relevant ideas. Intending and believing are the genuine intentional connections; the relation between the utterance token and the idea that comes to be its meaning is entirely derivative. Unlike intending and believing, it plays no role in the primary causal and phenomenal order of the world. Indeed, on a broadly Gricean approach, utterance-token meaning is explicitly definable in terms of these underlying intentional connections.<sup>28</sup>

The reader may ask how in a given case one can determine that a relation is ad hoc and not a genuine connection. There are two related procedures. First, we do, as a matter of fact, have quite firm intuitions about these matters generally. The following familiar criterion often helps to bring out these intuitions: predicates that are not just syntactically primitive but semantically primitive as well always express qualities and connections. Since traditional philosophical analysis strives to systematize our conceptual intuitions within an ordered framework of definitions, it can often settle questions regarding semantical primitiveness and, in turn, questions of whether or not a relation is basic. This sort of procedure was just illustrated in our application of a broadly Fregean philosophical analysis of numerical measure functions and in our application of a broadly Gricean philosophical analysis of utterance-token meaning. Second, this sort of procedure meshes with another procedure that one can follow to help settle problematic examples that arise in the context of empirical inquiry. This second procedure, which is built upon the first, is this. Given the special role that qualities and connections play in phenomenal description—and in the constitution of experience itself—we may look to our experience to identify certain genuine qualities and connections, namely, phenomenal qualities and mental connections. (For example, we can experience green but not grue. If we were unable to identify phenomenal qualities and mental connections in this way, we could not notice change or constancy in our experience.) Having identified these, we may then seek causal explanations of why we experience them when we do. Among the competing explanations, consider those that posit theoretical qualities and connections described solely in terms of known phenomenal qualities and mental connections, the notion of causation, the general notions of quality and connection, and any other basic notions isolated by our best system of philosophical analysis. Since the explanations are all formulated with the same terms, one can straightforwardly compare their complexity without running into the relativist's worry that ad hoc properties and relations might sneak in under the veil of a superficially simple syntax of primitive theoretical terms. After doing this, one would be justified in identifying the simplest of these explanations as correct. Then, from this explanation one can extract a provisional list of theoretical qualities and

connections. Suppose, however, that this procedure should fail to isolate a unique causal explanation—and, hence, a unique list of theoretical qualities and connections. The resulting situation would not be revolutionary; it would be just one more instance of the familiar problem of the underdetermination of theory by the data.

(2) *Infinitive and gerundive constructions*

Philosophers with an interest in intentionality have focused recently on those intentional phenomena that are naturally reported in language by means of 'that'-clause constructions: for example, 'x believes that A', 'x doubts that A', 'it appears to x that A'. Less attention has been paid to those intentional phenomena that are naturally reported in language by means of infinitive and gerundive constructions: for example, 'x intends to F', 'x decides to F', 'x wants to F', 'x fears F-ing', 'x imagines y F-ing', etc. We have seen that the proposed analysis fits nicely those intentional relations associated with 'that'-clause constructions, as in the relations of believing, doubting, being appeared to. For example, 'x believes that A' is parsed as:

x believes that A.

And the singular term 'that A' denotes an intensional entity belonging to the second traditional conception, for the occurrence of A in 'that A' has the type of fine-grained intensionality characteristic of that conception. Accordingly, 'believes' expresses a relation that characteristically holds between an individual and a conception 2 intension. Further, this relation holds between these items as a contingent fact independently of whether it holds between the individual and necessarily equivalent intensions. Thus, it fits the analysis. And the same thing goes for the relations of doubting, being appeared to, etc. But does the analysis fit those intentional relations associated with gerundive and infinitive constructions—the relations of intending, fearing, etc? The answer is that it does.

Let us suppose that these infinitive and gerundive constructions should be taken at face value syntactically. Then, for example, 'x intends to F' and 'x fears F-ing' would be parsed respectively as follows:

x intends to F.  
x fears F-ing.

Given this parsing, the singular terms 'to F' and 'F-ing' would in at least some cases denote intensional entities belonging to the second traditional conception, for in at least some cases the occurrences of F in these singular terms have the type of intensionality characteristic of the second conception. (For example, it is possible for someone to intend to build something whose top surface forms a trilateral and yet not to intend to build something whose top surface forms a triangle. And it is possible for someone to fear riding a

cycle with fewer than three wheels and yet not to fear riding a cycle with fewer than  $\sqrt[3]{27}$  wheels.) Accordingly, 'intend' and 'fear' would express relations that can contingently hold between an individual and a conception 2 intension and can do so independently of whether they hold between the individual and necessarily equivalent conception 2 intensions. Thus, if these infinitive and gerundive constructions are taken at face value, the associated intentional relations straightforwardly satisfy the proposed analysis. The same thing goes for the other relations in this family—deciding, wanting, etc.

It should be noted, incidentally, that each of the verbs 'intend', 'decide', 'want', 'fear', etc., can take 'that'-clauses as well as infinitives and gerundives: for example, 'x intends that x himself will *F*', 'x decides that x himself will *F*', 'x fears that x himself *Fs*'. Since such 'that'-clauses straightforwardly denote conception 2 intensions, the associated intentional relations—intending, deciding, fearing—satisfy the analysis quite independently of the matter of infinitive and gerundive constructions. Indeed, someone might try treating these infinitive and gerundive constructions not at face value but as transformations from certain underlying 'that'-clause constructions. For example, someone might try treating 'x intends to *F*' as a transformation from 'x intends that x himself will *F*'. If this works, the issue of infinitive and gerundive constructions would not even arise.

### (3) *Objectival constructions*

Constructions of the following kind are frequently used to report intentional phenomena: 'x looks for *y*', 'x wants *y*', 'x thinks of *y*', 'x is interested in *y*', 'x sees *y*', 'x loves *y*'. Yet on the surface, such constructions hardly seem to be stating that an individual stands in an intentional relation to a conception 2 intension.

There are two ways to deal with this family of examples. First, following Church, Quine, and others, we might treat objectival constructions as transformations from logically prior constructions that require right on their surface that an individual stand in an intentional relation to an intervening conception 2 intension. For example, 'x looks for *y*' and 'x wants *y*' might be treated as transformations from 'x strives to find *y*' and 'x wants to have *y*', respectively. And we have just seen that intentional connections associated with such infinitive constructions straightforwardly mesh with the proposed analysis of intentionality. The other objectival constructions can be dealt with in analogous ways.

Although the transformational treatment might seem a bit forced, there is good reason to take it seriously. Suppose we were instead to treat objectival constructions as ordinary relational formulas. Then, because a sentence like 'Ponce de Leon looks for the Fountain of Youth' seems true, we would seem to be forced to hold that Ponce de Leon actually stands in some relation to the Fountain of Youth. But since the Fountain of Youth does not exist, we

would be forced to hold, as Meinong did, that there literally are objects that do not exist, that there literally are unreal things. When fully generalized, this unrealism might be more than ontologically excessive; it might produce insuperable logical difficulties.<sup>29</sup> However, if we were to adopt the transformational approach, the offending occurrences of vacuous names and vacuous descriptions would give way to occurrences within intensional contexts, occurrences free of problematic ontological commitments. Indeed, even the thorniest instances of the problem of intentional identity could be resolved without recourse to nonexistent objects.<sup>30</sup>

The second way to deal with objectival constructions is to take them at face value. After all, each of the relations expressed by these objectival constructions either is identical to—or is necessarily included in—a relation that can contingently hold between an individual and an idea independently of whether it holds between the individual and necessarily equivalent ideas. For example, not only do individuals think of other individuals, but also they think of ideas; not only do individuals look for other individuals, they look for new ideas (new theorems, new concepts); not only do individuals want other individuals, they want new ideas (new strategies, new scenarios). And this is often so independently of whether they stand in these relations to necessarily equivalent ideas. Thus, if objectival constructions are to be taken at face value, the associated mental relations satisfy the proposed analysis straight off. In this case, however, I must add a bit more to the account of the aboutness of intentional phenomena.

According to that account, fine-grained intensions are the items that are in the primary sense about other objects. An intentional act is about objects only secondarily, inasmuch as it involves standing in an intentional relation to an intension that is about those objects. However, if objectival constructions are to be taken at face value, then a person will perform an intentional act if he stands in one of the associated objectival relations, not to an intervening intension, but directly to the object the act is about. For example, I perform an intentional act of looking for this pen if I stand in the looking-for relation, not to an intension that is about this pen, but to the pen itself. Would this show that the proposed account of the aboutness of intentional phenomena needs revision? Not clearly.

To see why, consider the following thesis: for each of the problematic objectival connections  $d$ , there is an intentional connection  $c$  such that, necessarily,  $d$  connects an individual  $x$  to an object  $y$  if and only if  $c$  connects  $x$  to an intension that is about  $y$ . (For example, necessarily,  $x$  looks for  $y$  if and only if  $x$  endeavors to find  $y$ ; endeavoring is an intentional connection, and that which  $x$  is endeavoring, namely, to find  $y$ , is an intension that is about  $y$ .) Suppose that this thesis is correct. It follows that the condition that  $d$  connects  $x$  to  $y$  is the same as the condition that  $c$  connects  $x$  to an intension about  $y$ . Therefore, since phenomena are just contingent conditions, the

phenomenon of  $x$ 's standing in relation  $d$  to  $y$  is just the phenomenon of  $x$ 's standing in relation  $c$  to an intension about  $y$ . Given this, my account of the aboutness of intentional phenomena remains intact even if objectival constructions are taken at face value.

Is the above thesis correct? It, or some variant of it, certainly seems to hold for most of the objectival relations in question. However, for the sake of argument let us explore the possibility that there are exceptions, perhaps the relations of loving and attending. Is it really true that someone can genuinely love something without, in the very act, loving to know it or loving to have it? And can someone genuinely attend to something without, in the very act, realizing that it participates in a figure/ground relation or that he himself is aware of it? These possibilities certainly may be doubted. However, it would be good for us to have waiting in the wings an account that does not turn on these issues. I will suggest one that is based upon a distinction between directedness and aboutness.

Let us suppose that the above possibilities of loving and attending are genuine, that it is truly possible to love an object or to attend to it without the indicated sort of conceptual mediation. In this case, although it would still be quite appropriate to say that the phenomenon of loving the object and the phenomenon of attending to the object are *directed toward* the object, we could not say that these phenomena are *about* it. Aboutness arises only via conceptual mediation; that is, aboutness enters just at the point when an individual stands in an appropriate intentional connection to a concept or thought that is about the object.

To spell out this account precisely, we will make use of the notion of a determinate intentional connection. Some intentional connections are species of others in the sense that they are necessarily included in them. For example, attending is a species of awareness because it is necessary that, if  $x$  attends to  $y$ , then  $x$  is aware of  $y$ . An intentional connection is *determinate* if and only if no other intentional connection is a species of it, i.e., no other intentional connection is necessarily included in it in this fashion. We need this notion for the following reason. Suppose that a traditional empiricist theory of pure, nonintentional experience is correct. For example, suppose that a pure experience of a fragrance or of undirected anxiety is possible. If so, such a phenomenon would be neither about anything nor directed toward anything; it merely occurs. Or so the traditional empiricist would have it. (On our analysis, the reason that the empiricist's relation of pure experience does not qualify as intentional is that thoughts and concepts are necessarily excluded from its range. See the next section for a fuller discussion of the empiricist's theory.) Notice that, necessarily, if  $x$  has a pure experience of  $y$ , then  $x$  is (in some sense) aware of  $y$ . In such a case, the phenomenon of  $x$ 's being aware of  $y$  (say, the feeling of undirected anxiety) would not be directed toward or be about anything. Now the relation of awareness satisfies our

analysis of intentional connection because (on other occasions) thoughts and concepts can occur hyperintensionally in its range. So merely standing in just any intentional connection (say, awareness) to an object is not sufficient for being directed toward the object. However, unlike general intentional connections like awareness, determinate intentional connections always guarantee directedness. (For example, if I am actually attending to the feeling of anxiety, the phenomenon of my doing so is of necessity directed toward this feeling.) Indeed, the restricted notion of a determinate intentional connection leads immediately to an analysis of directedness.

With this explanation in mind, I suggest the following as a first approximation: (elementary) phenomenon  $p$  is *directed toward* an object  $y$  if and only if, for some individual  $x$  and some determinate intentional connection  $c$ ,  $p$  is the phenomenon of  $x$ 's standing in relation  $c$  to  $y$ . But how, on this approach, does aboutness arise? The answer is very much in the spirit of our original analysis. Aboutness is a species of directedness, namely, one that involves conceptual mediation: (elementary) phenomenon  $p$  is *about* an object if and only if  $p$  is directed at a thought or concept that is about the object. Minor adjustments might be called for in this account of directedness and aboutness,<sup>31</sup> but it seems safe to say that a fully satisfactory analysis of intentional phenomena is at hand even if, as we have been doing, problematic objectival constructions are taken at face value.

#### (4) *Nonintentional relations to ideas*

There is a family of relations that can contingently relate individuals to ideas and that give an initial appearance of being basic. When we examine them, however, most or all lose that appearance. In any event, each of these relations can be disqualified as a counterexample on another ground: most of them do not appear even superficially to be hyperintensional. Consider the sentence ' $a$  is disposed to  $F$ '. Suppose we parse it this way:

$a$  is disposed to  $F$ .

And suppose the infinitive phrase ' $to F$ ' denotes the concept of being something  $x$  such that  $Fx$ . Then, the relation of being disposed would be able to relate contingently an individual to a concept. But this relation would not be a counterexample to our analysis, for clearly it is not hyperintensional. For example, it is necessary that, if one is disposed to be depressed, then one is disposed to be not not depressed. Likewise for any other infinitive phrase necessarily equivalent to ' $to F$ '. Other relations like being disposed include tending, approaching (as a limit), needing, and even the ought relation: a person tends to be depressed, an electron approaches (moving at) the speed of light, Sister Teresa resembles (being) an angel, the plant needs to have more water, the soup ought to have more salt. As with the relation of being disposed, each of these fails to be hyperintensional.<sup>32</sup>

There are, nevertheless, four kindred relations that some people might deem hyperintensional and that are not intentional connections: (1) a certain relation of efficient causation, (2) organismic striving, (3) final causation, and (4) natural meaning. (For example, it might be said that the car *causes* the boy to fall, the liquid *strives* to be in a state of equilibrium and the plant *strives* to be in the sunlight, the kidney *functions* [so as] to remove wastes from the blood, the red spots *mean* that the child has measles.) Because some or all of these relations might be hyperintensional, it would be good to be able to show on other grounds that they do not qualify as counterexamples. I will indicate briefly why most philosophers today would agree that there indeed are such grounds.

First, to most philosophers the above causal relation should be counted as ad hoc on the ground that, as a categorial fact, a causal connection can only connect phenomena (or events) to phenomena (events). That is, only phenomena (events) can be true efficient causes or effects; concrete particulars (e.g., cars), in contradistinction to phenomena involving them, can be neither efficient causes nor effects.<sup>33</sup> Second, most philosophers would doubt that there is any genuine nonintentional connection answering to our casual talk about striving: anything that literally strives for ends has a mind; all other uses of 'strive' are only metaphorical and do not express the sort of relation that serves to fix the fundamental order of the world. Third, concerning final causes, the leading view is that our use of 'function' divides into two distinct types—intentional and natural. The former type is reducible to more fundamental intentional notions such as desiring, intending, trying, etc; the latter type, to more basic, purely mechanistic notions (plus perhaps certain normative notions such as fitness, health, or well-being).<sup>34</sup>

Finally, if natural meaning does not just boil down to the relation of rendering probable, which we dealt with earlier, it seems in any case not to be basic. After all, it is highly doubtful that any genuinely basic relation of natural meaning can ever really hold between an individual just on its own and an idea. For an individual to mean naturally, a third parameter certainly is required—namely, (i) a person or group of persons with a system of background beliefs or (ii) a background theory or context, where these items are treated as systems of propositions considered in abstraction from the intentional fact that they are believed by the person or group of persons in question. However, if these suppressed third parameters are brought in, the resulting explicitly ternary relations mesh smoothly with our analysis. For the first ternary relation is intentional and, hence, is not even a candidate counterexample in the first place. Although the second ternary relation is nonintentional, it fails to be a counterexample on the grounds that it does not meet the contingency requirement: whenever it holds, it holds necessarily.<sup>35</sup> [For example, suppose that (a) these spots on the child mean (b) that the child has measles given (c) the proposition that, if the child has spots like

these, the child (in all probability) has measles. Then this ternary meaning relation holds among these three items necessarily.] Furthermore, these two ternary relations are nonbasic: (i) the relation of  $x$ 's naturally meaning  $y$  to person(s)  $z$  is a definable intentional relation more or less equivalent to the relation of  $x$ 's making  $y$  evident to  $z$ , and (ii) the relation of  $x$ 's naturally meaning  $y$  given proposition(s)  $z$  is a logical relation no doubt definable in terms of the notion of logical consequence together with other purely logical notions.

These and analogous considerations will, I hope, convince most people that these four relations pose no threat to our analysis. But if there is residual doubt, the analysis can easily be tightened up so as to rule them out explicitly.<sup>36</sup>

Since I have been unable to find better candidate counterexamples than the foregoing, I am inclined to hold that the analysis is indeed free of all counterexamples. At most, minor adjustments would be called for to deal with one controversial philosophical theory or another.

#### 4. NONINTENTIONAL MENTAL PHENOMENA

According to the first half of Brentano's thesis of intentionality, all and only mental phenomena are intentional. Is it really true that intentionality is the mark of the mental? The counterexample that springs to mind is that of pure, uninterpreted experience—pure sensation or pure inner feeling—as posited by traditional British empiricists. Any such experience would certainly be a mental phenomenon, but it would not be *about* or *directed toward* anything. Brentano and other intentionalists of course want to deny that there is any such thing as pure experience. However, Brentano puts forward the first half of his thesis as analytic, or at least as necessary. Therefore, this half of the thesis would be undermined if pure experience were *merely possible* for some beings or other, not necessarily human beings. In the face of this threat, we would be wise to have an analysis of the mental that is neutral with respect to the possibility of pure experience.

Just as in the case of judgment, so in the case of pure experience there are both relational and nonrelational theories. We saw that nonrelational theories of judgment ran into difficulties over the matter of generality. It is predictable, therefore, that nonrelational theories of experience (for example, so-called adverbial theories) also run into difficulty over this issue. Briefly, natural language has an apparatus for comparing experiences in infinitely many general ways: for example, 'the sense experience of creature  $u$  is qualitatively exactly like that of creature  $v$ ', 'the sense experience of  $u$  is qualitatively exactly like that of  $v$  except for some colors', 'the sense experience of  $u$  is qualitatively exactly like that of  $v$  except that their color spectra are

inverted', etc. Nonrelational theories are unable to capture the full expressive power of this idiom except by resorting to constructions that are at least implicitly relational, and the semantical properties of these constructions evidently cannot be spelled out without reintroducing our familiar relational constructions in the metalanguage. (See section 2 for corresponding difficulties in the adverbial theory's attempt to avoid the relational treatment of judgment.) Therefore, it is simpler and more natural to adopt the relational theory of pure experience right from the start.

If one adopts a relational theory, one encounters a striking parallelism between the relation of pure experience and the familiar intentional connections, except that now qualities and conditions play the role that concepts and thoughts played before. Specifically, the pure experience relation can contingently connect an individual to a quality or condition independently of whether it connects the individual to any necessarily equivalent concept or thought. For example, in "raw" sense experience I can be connected to a sensible quality (say, a taste or a smell) and I can do so even if I am connected to no concept that is necessarily equivalent to it. And in "raw" sense experience I can be connected to a condition (say, that something red is surrounded by blue) independently of whether I am connected to any necessarily equivalent thought.<sup>37</sup> Likewise for inner feelings: I can be connected to a pure feeling (say, the quality of sadness or pain or anxiety) independently of whether I am connected to any concept that is necessarily equivalent to the quality. Or so a theory of pure experience goes.

According to this account, then, the difference between pure and interpreted experience lies in the intensional objects: in pure experience the intensional objects are conception 1 intensions—qualities and conditions—whereas in interpreted experience the intensional objects are conception 2 intensions—concepts and thoughts. This suggests that pure experience and interpreted experience are mere modes or species of a single underlying relation of experience. Like pure experience, this general relation can contingently connect an individual to a quality or condition independently of whether it connects the individual to any necessarily equivalent idea. And like interpreted experience, the general relation can contingently connect an individual to a concept or thought independently of whether it connects the individual to any necessarily equivalent idea. But notice that the latter fact shows that this general relation satisfies the analysis of intentional connection.

On the picture that is emerging, pure experience is a species of this general relation, namely, the species with conception 2 intensions excluded from its range. And interpreted experience is a species of the general relation that arises via some other, perhaps rather more complex, operation (see the discussion of directedness and aboutness in section 3.3 for an indication of

what might be involved here). In either case, the species is necessarily included in the general relation: that is, necessarily, if  $x$  has a pure experience of  $y$ , then  $x$  experiences  $y$ , and if  $x$  has an interpreted experience of  $y$ , then  $x$  experiences  $y$ . This, together with the fact that the general relation satisfies the analysis of intentional connection, suggests that we may obtain an analysis of the full notion of a mental connection simply by extending the previous analysis so as to bring in all connections that are necessarily included in intentional connections.

We thus arrive at the following definition:

A connection is mental if and only if it is necessarily included in an intentional connection.

In symbols,  $c$  is a mental connection iff<sub>df</sub>

$(\exists d)(d \text{ is an intentional connection} \ \& \ \Box(\forall uv)(u, v \Delta c \rightarrow u, v \Delta d))$ .

All the intentional connections are of course mental according to the definition, for a connection is always necessarily included in itself as a limiting case.

With this definition in hand we may go on to the second stage of our approach. Using the definition of mental connection, we can define what a nonbasic mental relation is. Putting the two definitions together, we will have succeeded in characterizing what a mental relation is in general. Furthermore, the objections to this general analysis all seem to be variants of those facing the analysis of intentionality, and they can be disqualified for corresponding reasons.

On the view of the mind we now have, a descendant of the first half of Brentano's thesis of intentionality is vindicated, for intentional connections are the key to the analysis of mind. Nevertheless, we allow for the prospect of nonintentional experience as a special mode of experience. We are able to allow for this so easily just because the relation of pure experience is a species of the general relation of experience, which admits complex intensions into its range. In all cases, however, mental connections possess a contingent hyperintensionality, and this distinguishes them from all purely physical, naturalistic connections. Although beings without minds are connected to intensions in various ways, they can never be connected to them in this highly discriminating way. Indeed, the fact that no purely physical connection is mental may be viewed as the essence of the second half of Brentano's thesis of intentionality: given certain reasonable criteria for classifying phenomena, this fact implies Brentano's original claim that no purely physical phenomenon is mental. Suspended in a world of physical phenomena, the mind is a prism diffracting nature's stream of cause and effect into the colorful discriminations of thought and experience.

### 5. THE EXISTENCE OF MENTAL CONNECTIONS AND MIND

Our strategy for defining mental relations is two-staged. First, we say what it takes for a relation to be a mental connection. Then we use this notion to characterize the remaining, nonbasic mental relations: the relations whose definitions involve mental connections in some logically essential way. This strategy is based on the plausible premise that there are indeed mental connections. But critics might challenge this, alleging that all mental relations are nonbasic and, indeed, definable ultimately in terms of purely physical qualities and connections. According to these critics, the structure of the world would be fundamentally physical and formal.<sup>38</sup> In these closing pages, I will briefly sketch the ways in which I believe these critics can be answered.

One way to meet this criticism is phenomenological: regardless of its causal underpinnings, consciousness has a structure of its own that can be investigated phenomenologically, independently of the physical sciences. Such investigation reveals that familiar phenomenal properties are basic determinants of that structure and, analogously, that (at least some) familiar conscious relations are too. For example, the presence or absence in consciousness of nonbasic intensions can produce two types of detectable alteration: an alteration in the intentional contents of what one is thinking or a gestalt shift in how one is perceiving things. But unlike the presence or absence of such intensions, the presence or absence in consciousness of phenomenal properties can produce, not just these types of alteration, but a third type as well; namely, it can produce a shift in the "preinterpretive stuff" of consciousness. Phenomenal properties can thus be "in" consciousness in a particularly basic way. Similar considerations indicate that (at least some) familiar conscious relations can be "in" consciousness in an analogously basic way.

Another way to meet our critics' challenge is epistemological. There are two points to be made. First, let an ad hoc grue-like relation *experiencing\** be defined so that the following hold: *x* *experiences\** green iff *x* *experiences* green, and *x* *experiences\** blue iff (*x* has not just experienced green and *x* *experiences* blue now) or (*x* has just experienced green and *x* continues to *experience* green now). Assume what we suspect is false, namely, that the familiar relation of *experiencing* is on a par with this grue-like relation *experiencing\**. Then, '*x* *experiences\** green at  $t_1$  and *x* *experiences\** blue at  $t_2$ ' and '*x* *experiences* green at  $t_1$  and *x* *experiences* blue at  $t_2$ ' would be on a par too. But the change reported in the former sentence is a mere "Cambridge change"; only the second reports a real change. So *experiencing* is not on a par with the grue-like relation *experiencing\** after all. Generalizing on this, we are led to the conclusion that, if *experiencing* (and other conscious relations) were not among the genuine connections, there would be no plausi-

ble explanation of how it is that we successfully identify change or constancy in our conscious mental lives. The second epistemological point is this. The fact that we take phenomenal properties as qualities and the familiar mental relations as connections evidently plays a role in how ideally we should actually determine our list of theoretical qualities and connections in empirical science. (See section 3.1 for more on this procedure.) Failure to follow this procedure would seem to invite relativism on the matter of theoretical qualities and connections in science. But in this case, our physicalist critics might well lose the theoretical backing they thought they had for their thesis that no mental relation is a genuine connection and that physical qualities and connections suffice for the definition of all mental relations.

The third way to meet our critics is to refute straight off their physicalistic definability thesis. Physicalists have three strategies for attempting their definitions of the standard mental relations, strategies motivated by behaviorism, the mind-body identity thesis, or functionalism. When properly formulated (in some cases this requires considerable care), certain well-known arguments are, I believe, successful against these definitional strategies. I will briefly review some of them with an eye toward setting up a new argument, one purely logical in nature.

Behavioral definitions can be attacked by means of a "perfect pretender" argument: it is in principle possible for someone (to be disposed) to display the behavior typically associated with thinking  $p$  when, in fact, the person only pretends to think  $p$  out of a desire to deceive.<sup>39</sup> They can also be attacked by adapting W. V. O. Quine's argument for the thesis of the indeterminacy of radical translation: a speaker of a radically foreign language could share all our behavioral dispositions and yet think, e.g., that rabbit-hood is manifest on those occasions when we instead think that there exists a rabbit.<sup>40</sup> By either argument, it follows that behavioral dispositions do not pair up with a person's thoughts in the way required for behavioral definability of the thinking relation.

We come next to physicalistic definitions motivated by the mind-body identity thesis. These definitions can be attacked by a straightforward modal argument: even if it is a causal law that all and only creatures who think  $p$  have the neurophysiological property  $P$ , it is metaphysically possible for a being to think  $p$  and not to have  $P$  (or to have  $P$  and not to think  $p$ ); thus, the property of thinking  $p \neq P$ . And there is an epistemological argument: using nothing but introspection and pure reason, I cannot know that I have [neurophysiological property]  $P$ , yet using nothing but introspection and pure reason, I can know that I have the property of thinking something; so the property of thinking something  $\neq P$ .<sup>41</sup> Finally, there is the multiple-realization argument: the number of metaphysically possible physiological bases for thought is infinite, and it is impossible to give a finite specification in first-order physiological terms of what is common to them.<sup>42</sup>

As I indicated, I think there are sound formulations of all these arguments. However, suppose we come across some physicalists who simply deny some or all of the premises used in these arguments; for example, they might hold that, as a matter of *metaphysical necessity*, the mental supervenes on the physical (behavioral histories, physiological states, or even total physical worlds). And suppose, moreover, that they insist that there is no ontological significance in being able in principle to state a definition by finite means. For them, infinitary "definitions," though unstatable, are quite acceptable: if a mental relation is necessarily equivalent to, say, an infinite disjunction of physical properties and relations, this shows that the relation is physically definable and that it is not in a new category of nonphysical connections.

By careful argumentation we can, I think, refute the doctrines of modality and definition advocated by these physicalists. But for the sake of argument, suppose we were to allow these doctrines. We can construct a new, purely logical argument against the physicalistic conclusion. The key to our opponents' downfall is the failure to allow for the phenomenon of self-embedding. Consider the proposition that  $x$  thinks that everyone thinks many things. Is it the very same relation of thinking that "occurs" twice in this proposition? Contrary to ramified type theories, there are extremely compelling arguments that, in fact, it is the same relation. For example, how else are we to explain the important family of everyday "cross referential" dialogues like the following:

*A*: I think many things.

*B*: So do I; in fact, what you have just asserted is one of them [i.e., one of the things I think].

Here *A* asserts a proposition "involving" the relation of thinking, namely, the proposition that he thinks many things. Then *B* affirms the corresponding proposition about himself, namely, that he [i.e., *B*] thinks many things too. And then *B* goes on to provide an example of one of the things to which he stands in this relation of thinking, namely, the original proposition *A* asserted, which, as we saw, is a proposition "involving" this very relation of thinking. Everyday dialogues like this would make no sense if thinking could not "occur" in propositions that fall within its very own range.<sup>43</sup>

Now our opponents posit an infinitary definition of the form:

$x$  thinks  $y$  iff<sub>df</sub> ( $x$  has  $P_1$  &  $y =$  the proposition that  $Fa$ ) or ( $x$  has  $P_2$  &  $y =$  the proposition that  $Fb$ ) or . . . .

where *ex hypothesi*  $P_1$  is a physical property that is metaphysically necessary and sufficient for thinking the proposition that  $Fa$ ;  $P_2$ , for thinking the proposition that  $Fb$ ; etc. Here is the problem. What happens when we come to a proposition about thinking itself, such as the proposition that someone thinks something? Let it be granted that there is a physical property  $P_i$  that

is metaphysically necessary and sufficient for thinking this proposition. What does the associated disjunct in the infinitary definition look like? It cannot be this:

$x$  has  $P_i$  &  $y =$  the proposition that someone thinks something,

for then the definition would appeal explicitly to *the very relation being defined*. Alternatively, our opponents might try to offer the following candidate disjunct:

$x$  has  $P_i$  &  $y =$  the proposition that, for some  $u$  and  $v$ , either  $u$  has  $P_1$  &  $v =$  the proposition that  $Fa$  or  $u$  has  $P_2$  &  $v =$  the proposition that  $Fb$  or . . . or  $u$  has  $P_i$  &  $v =$  the proposition that someone *thinks* something or . . . .

In symbols,

$$P_i x \text{ \& } y = [(\exists uv)((P_1 u \text{ \& } v = [Fa]) \vee \dots \vee (P_i u \text{ \& } v = [(\exists wz)w \text{ thinks } z]) \vee \dots)].$$

But plainly this does no good, for the same problem just recurs: the psychological relation being defined has not been eliminated in favor of physical terms, but instead it occurs explicitly in the definiens one level down.

There are sophisticated logical techniques for avoiding this explicit circularity. However, if our opponents base their case on such techniques, it can be shown that the resulting overall physicalistic theory suffers new defects that make it unacceptable in comparison to a theory that treats the standard mental relations as connections.<sup>44</sup> So this hope for physicalistic definitions inspired by the mind-body identity thesis appears to be thwarted.

The only remaining strategy open to the physicalist is the functionalist one. The central premise of functionalism—and the premise upon which functional definitions are based—is that the standard mental relations are uniquely determined by their causal roles in functioning organisms. Therefore, the most direct way to undermine the prospect of functional definitions is to show that there exists a system of deviant relations that are demonstrably different from the standard mental relations and yet are causally and functionally indistinguishable from them. It turns out that we can construct just such a system of deviant relations by means of a certain generalization of a Quinean indeterminacy argument.<sup>45</sup> Thus, functional definitions are in vain, and the final strategy of the physicalist proves barren. It is important to note, moreover, that the possibility of this anti-functionalist argument is tied to the hyperintensionality of the standard intentional relations and, therefore, can be adapted to undercut the prospect of defining these relations in (naturalistic) information-theoretic terms.<sup>46</sup>

Ironically, nothing prohibits one from attempting a dualist theory of the standard mental relations that is akin to the old functionalism. Specifically,

one could revise the old functionalist definitions by explicitly requiring that the key relations over which the predicate variables range be genuine connections, not deviant relations. But if the resulting definitions were correct, they would have as a consequence the thesis of the present section, namely, that the standard mental relations are indeed genuine connections.

On the basis of the arguments reviewed in this section I believe that the overall definitional strategy of this paper is immune to physicalist attack. But there is one more question to ask. Must our theories posit mentality in the first place? According to eliminative materialism, the answer is negative: the standard mental notions are just unscientific; with the progress of science these notions will be seen to be on a par with those of alchemy and astrology. Even if this radical claim is granted for a moment, it would hardly show that the proposed analysis of intentionality and mind is without philosophical value, for in philosophy the value of an analysis does not ride on the utility a notion plays in science. Consider an analogy. Even if it turns out that we have no free will, philosophers will continue to be interested in defining the notion. For they will want to know what it is we lack, and indeed, they can use a precise analysis to improve their arguments that we truly are not free. Furthermore, even if eliminative materialism—at least as it is usually formulated—were true, that would not call into question the correctness of the proposed analysis of intentionality and mind. To be correct, an analysis is required to fit neither more nor less than the actual and hypothetical cases to which our mental notions are applicable. Eliminativism implies at most that there are no actual cases. But since our analysis takes no stand on the existence of actual cases, it is strictly independent of eliminativism. There would be trouble for the analysis only if it were impossible in principle for any mental relation ever to be realized in any being whatsoever. However, eliminativism hardly implies this highly counterintuitive thesis. In any event, this worry would be dissolved automatically if we could show the eliminativist that every acceptable comprehensive theory must posit actual cases of mentality. Let us in closing give a brief sketch of how we might try to show this.

The goal of theory is not restricted to the prediction and explanation of what we observe. Any acceptable comprehensive theory of the world (and indeed any acceptable comprehensive psychological theory) must also account for the fact that the theory itself has been arrived at by rational means, that it has epistemic merit—in short, that it is acceptable.<sup>47</sup> Because terms for acceptability, epistemic merit, rationality, etc., do not presently belong to physical theory, something like the following theoretical statement must be adjoined to the theory to obtain the requisite account:

A theory is acceptable for a person if and only if it is (or belongs to) the simplest overall theory that explains the person's data.

But as we saw in section 1, this statement exceeds the resources of nominalist and extensionalist physical theories. At a minimum, an apparatus for representing definitional relationships and for dealing with the notions of simplicity and explanation is required, and our rich intensional logic for qualities and concepts is (part of) the simplest theory that fills the bill. But so far this need only be Platonistic, not mentalistic. To discover the inevitable mentalist ingredient in any acceptable comprehensive theory, we must look to the matter of data.

After years of searching, philosophers have learned that there is no faintly plausible definition of the notion of data that does not directly or indirectly bring in some mentalist notion or other—observation, experience, pure experience, sensation, etc.<sup>48</sup> In the notion of data, therefore, we find the mentalist ingredient we are seeking.

Nevertheless, moderate eliminative materialists still have available an intermediate position that poses a threat to the full intentionalist picture of ourselves that we have been developing. According to this intermediate position, (1) all mental notions are eliminable except the notion of data; (2) the notion of data is definable in terms of pure experience without appeal to any intentional notions (for example,  $p$  is a datum for a person  $x$  iff<sub>df</sub>  $p =$  the proposition that  $x$  has a pure experience of  $y$ , where  $y$  is some sensible quality or condition, and  $p$  is true);<sup>49</sup> (3) the notion of pure experience is really materialistic because, unlike intentional notions, it has a first-order physiological definition.

Now I believe that the modal argument, the epistemological argument, and the multiple-realization argument can be formulated to show that physiological definitions of pure experience are either mistaken or unworkable for the epistemological purposes at hand. But even without resorting to these arguments, we can fault this moderate anti-intentionalist position. For there are two closely related vicious circles in which the resulting theory of acceptability is caught. The first we encountered in section 1: since our opponent's theory disallows logical and conceptual intuitions as data, how are the definitions in this theory to be justified? Why not adopt different definitions of data, explanation, and acceptability itself? Is the preference for the anti-intentionalist's definitions more than a bias or historical accident? It does the anti-intentionalist no good to appeal to the resulting simplicity of the overall theory, for there exist alternative definitions that yield even simpler, albeit *intuitively* unacceptable, overall theories. The second difficulty arises when one attempts to apply the anti-intentionalist theory in actual empirical situations. How can a theoretician  $x$  tell whether to accept into his overall theory a proposition  $p$  to the effect that he has a pure experience of some sensible quality or condition  $y$ ? On what basis does  $x$  accept propositions like  $p$  into his overall theory? It will do  $x$  no good to answer that  $p$  is data, for on what basis can he say that? According to his definition of data,  $p$  is data for  $x$  if

and only if  $p$  is an elementary proposition about  $x$ 's pure experience and  $p$  is true. So to show that  $p$  is data,  $x$  must already be able to show that  $p$  is true. No progress at all will have been made. At the same time, it does not help  $x$  to appeal to the simplicity of his overall theory, for its identity is fixed by the very propositions, like  $p$ , that he initially accepts into the theory as data. But that is what is in question: on what basis does  $x$  accept propositions like  $p$  into his theory in the first place? After all, by not accepting *any* propositions like  $p$ ,  $x$  could have a very simple theory indeed. It appears that  $x$  has no choice but boldly to accept  $p$  and dogmatically to decline to give any further justification for it. Now this move might be acceptable; however, if  $x$ 's theory of acceptability is comprehensive, it certainly must include the proposition that not just any proposition is acceptable straightaway without further justification. Given this,  $x$ 's theory, if comprehensive, must also include an explanation of what it is about propositions like  $p$  that makes them acceptable straightaway without further justification. However,  $x$ 's anti-intentionalist, purely physiological theory of data and sense experience seems quite incapable of providing such an explanation.

As far as I can tell, the only way out of these difficulties is to accept an explicitly intentionalist theory of acceptability. Such a theory might go as follows (with suitable qualifications): to accept a proposition or a theory is just to believe it; and for any elementary conceptual or logical proposition that is clearly and distinctly understood and for any elementary proposition about one's own present conscious states (including one's own conscious intentional states), one's mere belief in the proposition entails that the proposition is true. (Or more cautiously: it is necessary that *most* of the propositions like these that one believes must be true.) It is for this reason that such propositions are acceptable to the person without any further justification. If a theory of this general sort provides the only solution to the difficulties confronting the anti-intentionalist theory, then any acceptable comprehensive theory must posit intentionality at least in the theoretician himself.<sup>50</sup> And if the proposed analysis of intentionality is correct, such intentionality has a special logical status that necessarily distinguishes it from the physical. It would follow that any pure materialist is unable to reach the conclusion that his is an acceptable theory. An intentionalist can, but only by acknowledging the logically distinctive status of his own mind.

## Notes

1. I have dealt with this topic in lectures at several philosophy departments and also in chapter 10 of *Quality and Concept* (Oxford, 1982). The purpose of returning to the topic at this time is to give a revised analysis, together with an extended defense, that responds to the long list of provocative comments I have received. The following are only some of the people whose comments have helped me: Bruce Aune, Mark Bedau, Jose Benardete, Jonathan Bennett, Romane Clark, Crawford Elder, Kit Fine, Jaegwon Kim, Keith Lehrer, Michael Loux, Bill Lycan, Warren Quinn, David Smith, Chris Wagner, and Peter Woodruff. I am particularly grateful to Carol Voeller, whose probing comments on each draft were essential, and to George Myro, for numerous rewarding conversations on epistemic appraisal and acceptability. Finally, I thank the National Endowment for the Humanities for its generous support.

2. Here I am adapting an argument from George Myro's important paper "Aspects of Acceptability," *Pacific Philosophical Quarterly* 62 (1981): 107-17.

3. See chapter 1, *Quality and Concept*, and "New Foundations for Intensional Logic," *Handbook of Philosophical Logic* (Dordrecht, in press) for a detailed defense of this assertion.

4. See Nelson Goodman, *Fact, Fiction, and Forecast* (Cambridge, Mass., 1955), 75ff.

5. "A Formulation of the Logic of Sense and Denotation," *Structure, Method, and Meaning: Essays in Honor of Henry M. Scheffer*, edited by P. Henle, H. M. Kallen, and S. K. Langer (New York, 1951), 3-24.

6. *Ibid.*

7. For example, "Intensional Isomorphism and Identity of Belief," *Philosophical Studies* 5 (1954): 65-73.

8. See Alonzo Church, "On Carnap's Analysis of Statements of Assertion and Belief," *Analysis* 10 (1950): 97-99.

9. Hilary Putnam's "fundamental magnitudes," David Armstrong's "universals," Sydney Shoemaker's "properties," and David Lewis's "natural properties" are examples of qualities and connections. See Putnam, "On Properties," *Essays in Honor of Carl G. Hempel*, edited by N. Rescher, et al. (Dordrecht, 1970), 235-54; Armstrong, *Universals and Scientific Realism*, 2 vols. (Cambridge, 1978); Shoemaker, "Causality and Properties," *Time and Cause*, edited by P. van Inwagen (Dordrecht, 1980), 109-35; Lewis, "New Work for a Theory of Universals," *Australasian Journal of Philosophy* 61 (1983): 343-77.

10. Thoughts (i.e., 0-ary conception 2 intensions) appear to be what John Searle calls "propositional contents," and, taken together, thoughts and concepts appear to be what he calls "intentional contents" (see *Intentionality* [Cambridge, 1983]). Moreover, conditions (i.e., 0-ary conception 1 intensions) appear to be what he calls "conditions of satisfaction." However, Searle has not tried to treat these types of entities within a systematic logical theory, and he seems to think that the relation holding between a propositional content and its condition of satisfaction is an unanalyzable representationalist primitive. On the approach I advocate, this relation is just the relation of correspondence from the traditional correspondence theory of truth, and it is definable within the sort of intensional logic I have been sketching. Furthermore, Searle defines intentionality as aboutness or directedness, but he treats the latter notions as unanalyzable representationalist primitives belonging to what he calls "the circle of intentional concepts." On my approach, aboutness and directedness can be defined without circularity in terms provided by intensional logic. Finally, Searle invokes a basic distinction between intrinsic and derived intentionality, but no definition is offered. On my view, the intrinsic/derived distinction can be defined, but only within the framework of an intensional logic that is realistic about qualities and connections (i.e., our distinguished categories of conception 1 intensional entities).

11. In *Quality and Concept*, I do not state general definitions of the notions of fundamental logical operation, thought-building operation, etc.; I simply give lists that are adequate for my purposes there. However, general definitions are possible in the language  $\mathcal{L}_\omega$  with  $\Delta$ . Though

certain details are omitted, the following should make clear the overall strategy behind these definitions:

$v$  is a condition  $\text{iff}_{\text{df}} (\exists z)(\exists u)u \Delta z = v$ .

$v$  is a quality  $\text{iff}_{\text{df}} |u \Delta v|_u = v$ .

$v$  is an  $n$ -ary connection  $\text{iff}_{\text{df}} \langle u_1, \dots, u_n \rangle \Delta v|_{u_1 \dots u_n} = v$ .

$\Box A \text{ iff}_{\text{df}} |A| = ||A| = |A||$ .

$v$  is  $n$ -ary L-determinate  $\text{iff}_{\text{df}}$

$\Box (\forall u_1 \dots u_n) (\langle u_1, \dots, u_n \rangle \Delta v \rightarrow \Box \langle u_1, \dots, u_n \rangle \Delta v)$ .

$v$  is an  $n$ -ary fundamental logical operation  $\text{iff}_{\text{df}} v$  is an univocal  $n+1$ -ary L-determinate connection whose range consists exclusively of intensions.

$v$  is an  $n$ -ary thought-building operation  $\text{iff}_{\text{df}} v$  is an  $n$ -ary fundamental logical operation not having conditions, qualities, or connections in its range.

Notice that intensional abstracts  $[\dots]_{\alpha}$  are used in these definitions; however, we could get along without them by using instead one of our standard notions of singular predication, namely, the one according to which predicating one thing of another sometimes results in one of the original things itself (i.e.,  $\text{Pred}(u, v) = v$  for some  $u, v$ ). This of course is the standard condition-building operation of singular predication. Circularity does not result if we use this notion in our definitions, for the goal here is to define the *general* notions of fundamental logical operation, etc. None of these notions are used in the definitions; rather, a standard notion of singular predication is used.

Incidentally, some commentators have asked what it takes for a thought-building operation  $f$  to "correspond" to or to be "associated" with a condition-building operation  $g$ . The answer, which is given in *Quality and Concept* (note 18, p. 276), is that  $f$  and  $g$  must be *equivalent*, i.e.,  $(\forall \alpha \beta) (\alpha \Delta f(\beta) = \alpha \Delta g(\beta))$ . This answer makes it clear that there is no circularity in the definition given in the book (p. 196) of the key relation of correspondence that holds between conception 2 and conception 1 intensions and that provides the basis for the traditional correspondence theory of truth.

12. See chapters 8 and 9 in *Quality and Concept* for a partial survey of these theoretical tasks; see also David Lewis's elegant paper "New Work for a Theory of Universals" for a forceful defense of a kindred point of view.

The primary differences between Lewis's theory and mine lies not so much in the applications as in the foundations. Whereas Lewis posits possible individuals (both actual and nonactual) plus sets thereof, I prefer actual individuals, qualities, and connections plus logical combinations thereof. Beyond its inherent naturalness and ontological realism, this traditional Platonistic approach has several other advantages. For example, whereas the set-theoretical possibilist approach evidently must take the notion of a "natural property" as primitive, it is definable on the Platonistic approach. And because some relations (e.g., belief) can be "constituents" of propositions falling in their very own range, an apparently severe formal problem of ill-foundedness arises for the set-theoretical possibilist when trying to construct these relations out of sets and possible individuals. (See section 5 of this paper for discussion of an analogous ill-foundedness that produces difficulties for first-order physicalistic definitions of belief inspired by the mind-body identity thesis.) A unique feature of the logic for qualities and concepts is that the theory for these important "self-embeddable" relations is already worked out. Of course, what makes this possible is that these relations are taken as primitive entities rather than as set-theoretic constructs. Incidentally, to accommodate fine-grained intensionality, set-theoretical possibilism appears to have no alternative but to identify thoughts and concepts with ad hoc entities such as sequences of "intensions" (i.e., sets of sets of sets of possible individuals). However, it is highly unintuitive that such ad hoc entities *really* are the sort of thing we believe.

13. This way of phrasing the principle invites a kind of self-application, reminiscent of the self-application involved in the Montague-Kaplan paradox of the knower, that might lead to conflict with Gödel's second incompleteness theorem. (For an accessible and illuminating study of the Montague-Kaplan paradox, see C. Anthony Anderson's "The Paradox of the Knower," *Journal of Philosophy* 80 [1983]: 338-55.) We may assume that this sort of difficulty can be avoided by adapting some device or other developed for resolving the logical paradoxes.

14. See *Quality and Concept*, 180f., and David Lewis, "New Work," 167f.

15. See *Quality and Concept*, section 47, for a detailed defense of this approach to logical validity. See also note 11 above.

16. See David Lewis, "New Work," 365-70, for a nice defense of this thesis.

17. Although he uses a different terminology, David Armstrong gives an appealing analysis of what is in effect the notion of a causal connection. See *Universals*, vol. 2, 148-57, and *What Is a Law of Nature?* (Cambridge, 1983). I advocate a similar approach; on my account, however, phenomena (and not qualities) are typically the relata of causal connections. (A phenomenon is a contingent condition wherein either a single item has a nonlogical quality or a number of items are connected by a nonlogical connection.) Incidentally, despite the affinities between Armstrong's theory and mine, the only intensions Armstrong acknowledges are causal and formal qualities and connections, whereas for me there are further intensions: for example, those constituting the contents of experience and thinking (i.e., phenomenal qualities, mental connections, and the full range of concepts and thoughts). I believe that these further intensions are unavoidable in epistemology and the theory of mind.

18. For a detailed defense, see *Quality and Concept*, chap. 1, and "New Foundations for Intensional Logic." General sentences about intentional phenomena (e.g., "Self-aware people think of all and only things that they think that they think of"; "for all  $x$  and  $y$ , if  $x$  has  $y$ , it is possible that someone knows that  $x$  has  $y$ "; etc.) also appear to be a problem for the self-ascription theories of belief recently developed by Roderick Chisholm (*The First Person*, Minneapolis, Minn., 1981) and David Lewis ("Attitudes *De Dicto* and *De Se*," *Philosophical Review* 88 [1979]: 513-43). Section 39 in *Quality and Concept* suggests a solution to the Castaneda-Perry puzzles that prompted the theories of Chisholm and Lewis, a solution that leaves intact the usual apparatus for making general intentional statements.

19. Some people might try to avoid this conclusion by adopting a substitutional treatment of quantifiers and by treating 'that'-clauses as vacuous singular terms. There are significant problems with this maneuver, however. See *Quality and Concept*, 252, note 5, and "New Foundations for Intensional Logic."

20. See "New Foundations for Intensional Logic" for a more complete critique of the adverbial approach and other nonrelational approaches.

21. This strategy would be undercut if there simply were no intentional connections or if there were ad hoc intentional relations whose analyses depended in no logically essential way upon intentional connections. In section 5, I will argue that these worries are unfounded.

The analysis of intentional connection I will give here differs from that given in *Quality and Concept*. Although I believe that (some version of) the original analysis is correct, the present analysis is perhaps easier to defend. Of course, nothing would prevent one from combining the two analyses. This might be the safest strategy in the end; see note 33.

22. I borrow this term from Peter Woodruff, who uses it more broadly as a term for any form of conception 2 intensionality.

23. I discuss this relation largely for heuristic reasons, for it hardly seems to be a genuine connection, the sort of relation that fixes the logical, causal, or phenomenal order of the world. Indeed, it is doubtful that this relation can ever really hold just on its own. Individuals do not seem to belong to a category of things that can just on their own render thoughts probable; an entire condition or thought (perhaps involving an individual) is the only sort of thing that might entirely on its own be able to do so. By the way, I assume an objectivist doctrine of probability in the text. If probability is ultimately a subjectivist notion, then probability relations would

be defined (partly) in terms of intentional notions. They would not, however, be basic psychological connections. (See note 28 for the definition of the notion of a psychological connection.)

Incidentally, some philosophers (e.g., Peter Achinstein, Elliott Sober) believe that there is hyperintensionality in nature. To see why these views do not threaten my account of intentionality, see section 3.4 and note 33.

24. See *Word and Object* (Cambridge, Mass., 1960), section 12. It is no coincidence that Quine believes that "Brentano's thesis of the irreducibility of intentional idioms is of a piece with the thesis of indeterminacy of translation" (p. 221).

25. Notice that the analysis in no way restricts the possible range of the connection *c*. As far as the analysis is concerned, items from any metaphysical category can be in *c*'s range. Notice also that the analysis is given for binary connections only; however, it can be generalized. Finally, for expository purposes I have stated the hyperintensionality condition in its simplest and weakest form. There are stronger versions available. For example, ' $y \approx_N z$ ' might be replaced with ' $x$  and  $y$  are analytically equivalent' or ' $y$  and  $z$  are relevantly equivalent'. Or we might use the following substitute analysis:

$$\Diamond(\exists xy)(\text{Ind}(x) \ \& \ \text{Idea}(y) \ \& \ (\forall z)(y \neq z \ \& \ y \approx_N z) \rightarrow \Diamond(x, y \Delta c \ \& \ x, z \nabla c)) \ \& \ \Diamond x, y \nabla c.$$

Further alternatives also come to mind. Nevertheless, the ensuing discussion will, I believe, show that the present version suffices.

26. These definitions of physical measure functions might also need to appeal to relations that hold between individuals and certain abstract entities known as *physical amounts* (1 ounce, 5 liters, 2 feet, etc.). Physical amounts, however, belong to the Aristotelian category of quantity and are not concepts at all. They certainly should not be confused with number concepts as the following equations make clear: 2 ounces = 56.7 grams and the concept of being (or weighing) 2 ounces  $\neq$  the concept of being (or weighing) 56.7 grams. (See C. D. C. Reeve, *Mass, Quantity, and Amount* [unpublished dissertation, Cornell University, 1980], for a careful discussion of these distinctions.) Now the primary relation(s) holding between an individual and its amounts might be purely physical (the piece of gold *weighs* 3 grams); on the other hand, the primary relation here might be purely logical, namely, a relation of predication (e.g., the piece of gold *is* 3 grams). Either way, measure functions that correlate individuals with number concepts (such as the relation holding between a physical thing and the number of grams in it) still would be plainly derivative.

I should mention that there is a rather different approach to the theory of measurement than the one sketched in the text. According to this approach, the derivative character of measure functions is explained by means of various representation and uniqueness theorems: given a sufficiently rich characterization of physical objects in terms of physical relations among them (e.g., in terms of equality of weight, equality of length), one can actually derive the existence of the usual numerical measure functions. Advocates of this approach think of these functions as nonbasic just because they are "extrinsic" to the underlying "intrinsic" physical relations that ensure their existence and fix their identity (up to multiplication by a positive real number). Whether or not this explanation is fully correct, the approach does draw attention to a feature of numerical measure functions that clearly is ad hoc. None of the functions in the infinite class of functions equivalent up to multiplication by a positive real number is more basic than any other; the decision to use one over another depends on an arbitrary choice of standard unit. For an elegant philosophical introduction to, and extension of, this approach to the theory of measurement, see Hartry Field's *Science without Numbers* (Princeton, N.J., 1980).

27. Anyone still in doubt about this conclusion could always add to the proposed analysis the further requirement that intentional connections must possess independent veracity, that is, that they can contingently connect some individual to some thought or concept independently of whether or not the thought is true or the concept has any instances. This move would explicitly block numerical measure functions as counterexamples, for a number concept must

have instances if it is the value of a measure function. (On some analyses, zero is a concept having no instances; however, it would be impossible for a thing to be related to zero with independent veracity since, on these analyses, zero *necessarily* has no instances.)

28. The relation of speaker meaning—i.e., the relation holding between a speaker  $x$  and an idea  $y$  such that  $x$  means  $y$  by uttering something—would not be a counterexample to our analysis since it is an intentional relation; of course, on a broadly Gricean analysis, speaker meaning is definable in terms of more basic intentional relations—intending, believing, etc. Now, like speaker meaning, utterance-token meaning fails to be a counterexample to our analysis simply on the grounds that it too is an intentional relation. But if, contrary to what I say in the text, utterance-token meaning were truly a connection, we should not be happy to leave things here. A goal of the analysis is to isolate those connections simply in virtue of which a creature (as opposed to an utterance token produced by a creature) is intentional, and utterance-token meaning would not be such a connection. Thus, if we were to agree that utterance-token meaning is a connection, we should want to implement the following routine. A plurality of intentional connections is deemed *minimal* if and only if all intentional connections can be defined in terms of them, where no smaller number of them suffices for this purpose. Then, an intentional connection is defined to be *psychological* if and only if it is necessarily included in the union of the connections in such a minimal plurality. These psychological connections are those in virtue of which a creature is intentional. Because utterance-token meaning is not psychological in this sense, this routine would solve our problem. However, I do not implement this routine, for, after all, utterance-token meaning is disqualified as a connection in the first place.

When I speak of utterance-token meaning in the text, I am referring to meaning relations that hold contingently between utterance tokens and intensions in virtue of appropriate utterers' intentions. What is the relationship between these intentional meaning relations and the abstract meaning functions that arise in formal semantics? The answer is roughly this. To establish a given language as their own, the members of a speech community institute a convention to produce tokens of utterance types in the language only if the meaning (as specified by an abstract meaning function) of the utterance type is the intension that the speaker (intentionally) means by producing the token. Because an abstract meaning function is an antecedently given, purely abstract pairing of utterance types and intensions, it does not qualify as a counterexample to the analysis of intentionality. First, these functions pair utterance types (universals) and intensions, but to be a counterexample they would have to pair particulars and intensions. Secondly, these functions violate the contingency requirement since they are L-determinate. Thirdly, because these functions are arbitrary pairings of utterance types and intensions, there is no reason to think that they are genuine connections.

29. For an indication of some of the logical difficulties that might arise, see my review of Terence Parsons's *Nonexistent Objects* (*Journal of Symbolic Logic* 49 [1984]: 652-55).

30. See, for example, the technique sketched at the end of section 39, *Quality and Concept*.

31. Here is an illustration. It seems possible for someone to be thinking of a concept and not to be thinking of any object that the concept is about. For example, since I have a special theoretical interest in singular concepts, perhaps I am able to think of the concept of being identical to the color red without thinking of the color red itself. If this is possible, it would seem odd to say that this phenomenon of my thinking of the concept is *about* the color itself. Thus, if we assume that thinking of is a determinate intentional connection (of course, this is doubtful), then we should want to adjust the analysis of aboutness just given in the text. The following would suffice: (elementary) phenomenon  $p$  is *about* an object  $y$  iff  $p$  is directed toward a thought or concept that is about  $y$  and the occurrence of this phenomenon entails the occurrence of an associated phenomenon that is directed toward the object  $y$  itself (that is, for some individual  $x$ , some determinate intentional connections  $c$  and  $d$ , and some thought or concept  $z$  that is about  $y$ ,  $p$  is the phenomenon that  $x$  stands in relation  $c$  to  $z$  and, necessarily, if  $x$  stands in relation  $c$  to  $z$  and  $z$  is about  $y$ , then  $x$  stands in  $d$  to  $y$ ). This adjusted analysis solves the above

problem as follows. *Ex hypothesi* my thinking of the concept of being identical to the color red does not entail that I am thinking of the color red itself. Likewise for any other determinate intentional connection: my thinking of the concept does not on its own entail that I am connected by such a connection to the color itself. Therefore, according to the analysis, merely standing in the thinking-of relation to the concept does not qualify as a phenomenon that is about the color. And this is the outcome we are seeking. At the same time, other phenomena that are actually about objects are correctly counted as such by the analysis. For example, the phenomenon of my doubting that this (i.e., the color red) is a dispositional property is about this color, and our analysis counts it as such: the proposition that this is a dispositional property is about this color; my doubting that this is a dispositional property entails that I am thinking of this color, and the phenomenon of my thinking of this color is directed toward this color.

32. Although these relations do not pose a threat to our analysis, they do bring out some interesting points of difference between intentional and nonintentional relations. Recall the relation of falling under a concept, which we examined earlier. Whenever this relation holds between an individual and a concept, the concept *must* always have *that very individual* as an instance. Let us call this feature *dependent instantiation*. Intentional connections typically do not have dependent instantiations: if an intentional connection can connect an individual to a concept, typically it can do so independently of whether the concept has that individual (or indeed any individual) as an instance. Now unlike the relation of falling under a concept—and like intentional connections—the relations just discussed in the text do not have dependent instantiations. Nevertheless, there is an important logical affinity these relations have to the relation of falling under a concept, an affinity that intentional connections all lack. Take the ought relation as an example, and consider how we treat the modals 'must' and 'might'. 'A body must have a location' is best treated as a transformation from 'It is necessary that a body have a location', which ascribes the modal property necessity to the proposition that a body has a location; likewise, 'Socrates might philosophize until dawn' is best treated as a transformation from 'It is possible that Socrates will philosophize until dawn', which ascribes the modal property possibility to the proposition that Socrates will philosophize until dawn. Uniformity would seem to demand that we treat the modal 'ought' analogously. If so, '*a* ought to *F*' would be a transformation from 'It ought to be that *Fa*', which ascribes the modal property of what ought to be (i.e., what is good) to the proposition that *Fa*. If so, it would follow that the proposition that *a* ought to *F* is necessarily equivalent to the proposition that it ought to be that *a* falls under the concept of being *F*. Generalizing, we obtain the following necessary equivalence: *x* ought *y* iff it ought to be that *x* falls under *y*. Put metaphorically, ought is a relation that acquires a dependent instantiation *in the ideal*, as the world approaches what ought to be (what is good). Indeed, each relation in the cluster just discussed in the text has this special logical feature: it acquires a dependent instantiation in a relevant modal ideal. Intentional connections are never like this.

This observation is rather vague, and there is more than one way to make it more precise. Here is one. We saw that ought is a relation *R* that necessarily satisfied the following scheme for a basic modal property *Q*:  $xRy$  iff  $Q[x$  falls under  $y]$ . And tending seems like this too; just let *Q* be probability or likelihood. After all, it is necessary that *x* tends to *F* iff it is probable (likely) that *Fx*. Needing is not exactly like ought and tending in this respect, but it is similar. For it necessarily satisfies the following scheme for a basic modal property *Q* and a basic normative property *G*:  $xRy$  iff  $Q[Gx \rightarrow x$  falls under  $y]$ . Just let *Q* be natural necessity and *G* be well-being or flourishing; necessarily, *x* needs to *F* iff it is a natural necessity that *x* has well-being (flourishes) only if *Fx*. And the other nonintentional relations in our cluster all satisfy these or similar modal schemes. Intentional connections, by contrast, are never tied to underlying basic modal properties in any of these ways.

33. This categorial fact also protects our analysis of intentionality from the worry voiced in the closing section of Elliott Sober's "Why Logically Equivalent Predicates May Pick Out Different Properties," *American Philosophical Quarterly* 19 (1982): 183-89. In this article Sober

constructs an example of a triangle-selecting device whose description allegedly involves fine-grained intensionality. Then in his closing section Sober asks whether examples like this might not threaten analyses of intentionality that depend upon fine-grained intensionality. The answer is that such examples do not threaten our analysis: intentional connections must be able to connect individuals to fine-grained intensions; causal connections, by contrast, can hold only between phenomena and phenomena (or between events and events). Indeed, Sober himself seems tacitly to accept this categorial fact, for he writes as though "a's having *F* causes b's having *G*" is the canonical form for statements of singular causation. Suppose, however, that someone were to deny this categorial fact and to insist instead that particulars really can cause fine-grained intensions or that fine-grained intensions really can cause particulars. And suppose, moreover, that this person alleges that the envisaged causal connections are hyperintensional. Now we could mount persuasive counterarguments. But we could also deal with this person simply by tightening our analysis. For example, we could require intentional connections to have not only contingency and hyperintensionality but also independent veracity. (The latter notion is characterized in note 27 above and in section 48 of *Quality and Concept*.) Then the envisaged causal connections would not qualify as intentional because they do not have independent veracity. After all, the car does not cause the boy to fall unless the boy *in fact* falls, and triangularity does not cause Sober's triangle-selecting device to select an object unless something *actually* is triangular.

34. Incidentally, according to an Aristotelian theory of final causes, if the function of, say, a kidney is to remove wastes from the blood, then, necessarily, this is the function of the kidney. Hence, this Aristotelian notion of function would violate the contingency requirement and, therefore, would not be an intentional connection according to our analysis. This is the desired outcome because the Aristotelian notion is not intentional.

35. The meaning relation envisaged by John Perry and Jon Barwise in their system of situation semantics is a somewhat similar ternary relation holding between a particular token, a meaning, and a system of constraints. This relation would not be a counterexample to our analysis, however, for in their system, whenever this relation holds, evidently it holds necessarily, hence violating the contingency requirement. If not, one could disqualify this relation as a counterexample by suitably adapting the strategy sketched for utterance-token meaning in note 28.

36. The following revised analysis would suffice for this purpose:

A connection is intentional iff it is identical to—or it is necessarily included in—a connection that can contingently connect an individual to an L-determinate idea independently of whether it connects the individual to any necessarily equivalent idea.

None of the problematic relations under consideration in the text can connect an individual to any L-determinate idea. Incidentally, since the consciousness relation satisfies this analysis and since the relations of pure and interpreted experience are necessarily included in the consciousness relation, this analysis could be adopted for use in the full analysis of mental connection given in the next section.

37. On the standard nominalist relational theory of sense experience, the range of the pure experience relation does not include sensible qualities and sensible conditions, but instead it includes mental particulars (*sensa*, sense data, phantasms, subjective spaces, etc.). However, no ontological economy is gained by adopting this theory, for such private mental particulars cannot be compared with one another in all the ways we ordinarily compare sense experiences except by invoking the full apparatus of sensible qualities at some level or other. So one is ontologically better off adopting the Platonistic relational theory described in the text. In any case, the definition of mental connection offered in the text is designed to handle the pure experience relation regardless which of these two theories is correct; for as far as the definition is concerned, items from any metaphysical category could be in the range of any mental connection. As a result, the analysis is compatible with the nominalist theory that every object of pure experience is a private mental particular.

38. For example, this is David Armstrong's position in *Universals and Scientific Realism*.

39. See, for example, Hilary Putnam, "Brains and Behavior," *Analytical Philosophy*, 2nd ser., edited by R. J. Butler (Oxford, 1963), 1-19.

40. See *Word and Object*, chapter 2 and section 45.

41. See, for example, Frank Jackson, "Epiphenomenal Qualia," *Philosophical Quarterly* 32 (1982): 127-36; Thomas Nagel, "What Is It Like to Be a Bat?" *Philosophical Review* 83 (1974): 435-50; Richard Warner, "In Defense of Dualism," unpublished manuscript. See also Saul Kripke, "Naming and Necessity," *Semantics of Natural Language*, edited by Gilbert Harman and Donald Davidson (Dordrecht, 1972), 253-355, 763-69, especially 335.

42. See, for example, Hilary Putnam, "Psychological Predicates," *Art, Mind and Religion: Proceedings of the 1965 Oberlin Colloquium in Philosophy*, edited by W. H. Capitan and D. D. Merrill (Detroit, 1967), 37-48.

43. See "New Foundations for Intensional Logic" for an extended defense of the thesis that the standard psychological relations are self-embeddable. A wealth of further arguments can be extrapolated from those Saul Kripke gives against Tarski's infinite hierarchy of distinct truth concepts for English; see "Outline of a Theory of Truth," *Journal of Philosophy* 72 (1975): 690-716, sections 1 and 2. A principal aim of *Quality and Concept* is to provide an intensional logic able to treat self-embeddable psychological relations without invoking the implausible, artificial distinctions posited in ramified type theories.

44. I give the argument in full in "A Disproof of the Mind-Body Identity Thesis from Self-embedded Attitudes" (forthcoming).

45. See my "Mind and Anti-Mind: Why Thinking Has No Functional Definition," *Midwest Studies in Philosophy* 9 (1984): 283-328. It is worth noting that the argument of this paper is not tied to the use of Quinean transformations that concern ontology (e.g., the universal/particular transformation); elementary logical transformations also suffice (e.g., a transformation that maps fine-grained propositions  $[(A \& B) \vee C] \& D$  to fine-grained propositions  $[D \& (C \vee (B \& A))]$ ).

Incidentally, functionalism also runs into fatal difficulties regarding self-embeddable attitudes; see "A Disproof of the Mind-Body Identity Thesis."

46. This result thus contradicts the central thesis of Fred Dretske's book, *Knowledge and the Flow of Information* (Cambridge, Mass., 1981).

47. Here again I am adapting the line of argument developed by George Myro, "Aspects of Acceptability."

48. For example, the eliminative materialist Paul Churchland is unable to state his account of scientific progress without making repeated use of the mentalist terms 'sensation' and 'intrinsic qualitative identity of . . . sensations'. See chapter 2, *Scientific Realism and the Plasticity of Mind* (Cambridge, 1979).

49. A moderate eliminativist might initially be drawn to the following alternate definition:  
 $p$  is a datum for  $x$  iff<sub>df</sub>  $p$  = the proposition that  $x$  has a pure experience of  $y$ , where  $y$  is some sensible quality or condition, and  $x$  accepts  $p$ .

The problem with this definition for the moderate eliminativist is that the term 'accept' is *prima facie* intentional. The advantage of the definition in the text is that it avoids intentional terms.

50. Because such a theory must appeal to self-embedded intentional relations (acceptance, belief), we have a transcendental justification for the starting point of our purely logical argument (sketched earlier) against the possibility of physicalistic definitions of intentional relations. With this in mind, we might consider building self-embeddability right into our analysis of intentionality.