

BARRY LOEWER

FROM INFORMATION TO INTENTIONALITY

Arabella believes that her cat, Glendower, wants to go out. Her belief has representational and semantic features. It is about Glendower; it represents him as wanting to go out, and it has truth conditions. Her belief also has causal and rationalizing powers. She opens the door *because* she believes Glendower wants to go out. If we think that Arabella and other believers are physical entities we are led to wonder how it is possible for a physical thing, whether it is composed of cells or micro chips, to have beliefs, desires and other propositional attitudes. This is the problem of intentionality. It has proved to be a very difficult problem. The source of the difficulty is that intentional and semantic concepts, reference, truth conditions, meaning etc. make no appearance in biological or physical theory. Additionally, beliefs have a normative dimension. They are assessable as correct or incorrect, rational or irrational. But the descriptions which occur in physical theory apparently are nonnormative. How can states which not only can represent but also misrepresent be captured in physical theory? The challenge for a philosopher who holds that intentionality is part of the natural order is just this: To show how it can be that certain physical states are capable of representation and misrepresentation and then to show how such states enter into the causation of behavior.

Behaviorism and the identity theory, each in its own way, identify belief (and other intentional states and processes) with something physical: the first with dispositions to behave and the second with states of the central nervous system.¹ There are numerous well-known problems with both of these *isms*, but setting them aside, it is interesting to note how little effort their proponents (and opponents) have devoted to showing that such physical states have representational capacities. It is one thing to be told that A's belief that it is snowing is identical to such and such a neural state. It is quite another thing to explain how it is that this neural state is able to represent snow.

Functionalism may seem to come closer to providing a physicalistic account of representation. According to some versions of func-

tionalism a belief is a state which can be characterized in terms of its computational role. Since computations involve the processing of symbols this version of functionalism explicitly recognizes beliefs as representational. But, whatever the view that beliefs have computational roles contributes to our understanding of cognition, it presupposes rather than explains their representational powers. The well-known difficulty is that syntactical and computational relations, no matter how complex, cannot by themselves characterize representational relations.²

Recently Dretske, Fodor, and Stalnaker have suggested an answer to the question of what makes the belief that *p* the belief *that p*.³ These authors agree that being a belief state can be functionally characterized (although they favor different functional characterizations). Their novel idea is the suggestion that what makes a particular belief state the belief *that p* is to be explained in terms of informational content. A rough general formulation of the approach is:

- (IS) A believes that *p* iff *A* is in a state *B*n* such that
- (a) *B*n* satisfies certain functional conditions and
 - (b) Given that conditions *C_n* obtain *B*n* carries the information that *p*.

I will call these accounts "informational semantics". They differ primarily in how "information" is understood and in the conditions *C_n* under which informational content coincides with belief content.

Before turning to details, I want to make a few observations concerning the nature of the project. (IS) is intended as providing a *reduction* of facts about what *A* believes to physical facts, similar to, e.g., "Water is H₂O".⁴ In view of this, (IS) must satisfy two adequacy conditions if it is to be a correct reduction. (1) (IS) and its instances must be law-like and true. The test of (IS)'s correctness is that the beliefs that it attributes must tally with those attributed by folk psychology. Its law-likeness would presumably come from its systematically associating belief states with their contents. (2) The states *B*n*, the conditions *C_n*, and the notion of information employed must all be specifiable without appeal to semantic or intentional notions. This means, for example, that the *C_n* are characterized without recourse to other beliefs, to meanings, and so forth. Although (IS) cannot presuppose intentional notions it can employ intensional ones.

There would be nothing wrong with using teleological concepts or counterfactuals in (*IS*) even though these are intensional.

Although information-theoretic theories of meaning are in initial stages of development and none of their proponents would claim to have provided a fully satisfactory account of the semantic properties of intentional states, it will still be valuable to examine them to assess their prospects for success. In this paper I will focus primarily on the accounts offered by Dretske and Fodor. I will argue that their accounts do not work. They can satisfy one of the adequacy conditions only at the expense of the other.

DRETSKE'S ACCOUNT IN KNOWLEDGE AND THE FLOW OF INFORMATION

Dretske's information-theoretic account of belief is developed in his book *Knowledge and the Flow of Information* (KFI). He says there that his aim is to show how to bake a mental cake using only physical yeast and flour. The key ingredient in his recipe is information. Dretske's basic idea is that one event carries information about another event in virtue of there being certain lawful relations between the events or between the types which the events exemplify. For example, there being red spots on a child's face carries the information that the child has measles since there is a lawful regularity between red spots and measles. He defines the information carried by a signal *r* as follows:

A signal *r* carries the information that *s* is *F* = The conditional probability of *s*'s being *F*, given *r* is 1. (p. 65)

As it stands Dretske's definition is not well-formed since "*r*" is a singular term referring to a signal but the condition in a conditional probability must be a proposition or event type. The obvious modification is that it is the fact that a signal has a certain feature that carries information. The definition would then be

r's being *G* (hereafter *G*(*r*)) carries the information that *s* is *F* iff $P(F(s)/G(r)) = 1$.

It is important to note that *G*(*r*) may carry one piece of information while *G**(*r*) may carry different information. For example, the frequency of a sound may carry the information that the British are

coming by sea while its being loud may carry the information that there are hundreds of them.

There are a number of features of Dretske's account of information which merit discussion. One concerns the notion of probability that occurs in it. How does Dretske intend it to be understood? None of the usual interpretations of probability seems capable of serving his purposes.⁵ In a footnote he offers the following explanation of conditional probability:

In saying that the conditional probability (given r) of s 's being F is 1, I mean to be saying that there is a nomic (lawful) regularity between these event types, a regularity which nomically precludes r 's occurrence when s is not F . (p. 245)

Using this account of conditional probability in the definition of information makes the information carried by $G(r)$ a relation between event types. An event of type G carries the information that there occurred an event of type F .

There is an obvious problem with the new characterization of information. A voltmeter's reading 7 volts (one of Dretske's examples) will not carry the information that the voltage in the wire is 7 volts since there is no law which precludes the voltage in the wire from being other than 7 volts when the meter reads "7 volts". Dretske's reply (ch. 5) is that the nomic regularity in virtue of which one event carries information about another is to be understood as relative to what he calls "channel conditions". These are standing matters of fact relative to which the information carried by a signal is determined. In the voltmeter example the channel conditions presumably are that the voltmeter is in working order, there are no strong electromagnetic forces deflecting the voltmeter's pointer, and so on. The characterization of information then becomes:

The occurrence of an event type G carries the information that an event of type F occurs iff it is a lawful regularity that when an event of type G occurs and channel conditions CC obtain then an event of type F occurs or has occurred, but it is not a lawful regularity that whenever CC obtains an event of type F occurs.

It is not clear even with this revision that the voltmeter's registering "7 volts" carries the information that there are 7 volts in the wire. Suppose that the voltage in the wire is determined by the strength of a

magnet and certain other prior conditions. Call these conditions C . Given C it follows by law that the voltage in the wire is 7 volts. But this means that if C is included in the channel conditions the voltmeter's registering "7 volts" does not carry the information that the voltage in the wire is 7 after all since it is not required, given the channel conditions, to preclude the occurrence of anything other than a voltage of 7 in the wire. Here is a related difficulty. Suppose that there are two voltmeters, V_1 and V_2 , attached to the wire and both register "7 volts". If we include the fact that V_2 registers "7 volts" in the channel conditions for V_1 , then V_1 's registering "7 volts" will not carry the information that the voltage in the wire is 7 volts. The same holds for V_2 . These examples show that the information carried by $G(r)$ is quite sensitive to what counts as channel conditions. Unfortunately, Dretske offers no general account of how channel conditions are determined.

Since Dretske's project is to provide a physicalistic reduction of belief and he uses information in the reduction, one might worry that the notion of channel conditions will sneak intentionality in through the back door, rendering his "reduction" circular. In fact he does make some surprising remarks concerning channel conditions. He says that whether or not C is included in the channel conditions for some $G(r)$ "is a question of degree, a question about which people (given their different interests and purposes) can reasonably disagree, a question that may not have an objectively correct answer" (p. 133). If what counts as channel conditions is relative to interests and purposes, then the information carried by a signal is likewise relative. He seems to be admitting that intentional notions are involved in the characterization of information. I do not know whether this relativity to interests undermines Dretske's attempt to construct a physicalistic account of intentionality in terms of information. Perhaps the relativity is similar to the apparent relativity of "world-similarity" to context in possible world semantics for counterfactuals. In this case the relativity does not show that the analysis of counterfactuals involves intentional notions. However, it is clear that it would be fatal to Dretske's project if in specifying the information carried by r 's being G the property G or the channel conditions are themselves characterized in intentional or semantic terms. That is, it would clearly be circular if Dretske was forced to explain the content of A 's beliefs in terms of the information, say, carried by A 's believing that it is snowing or that

information was determined by including in the channel conditions that *A* has certain beliefs or speaks a certain language.

In view of the previous discussion I think it is clear that Dretske owes us an account of "channel conditions" before we can say that he has clearly characterized "information" or that his characterization is free from semantic and intentional presuppositions. However, I will set my worries about channel conditions aside for now and proceed to describe his construction of belief content.

Dretske takes seriously the idea that cognitive agents are information processors. Certain states of the brains of higher organisms carry information about the organism's environment. They also guide its behavior in ways that are appropriate, given its needs and desires, to the information represented. For example, it is plausible that the tokening of a certain state in a frog's brain carries the information that a fly is buzzing in front of the frog and that this state plays a role in causing the frog to behave in certain ways, e.g., attempting to capture the fly, which are appropriate to its informational content. Perhaps in a similar way structures in human brains carry information about various aspects of the environment. Dretske's proposal is that beliefs are just such states of neural structures, states which carry information and which also play an appropriate role in guiding behavior.

However, it is not obvious that belief states can be constructed out of information bearing states. One problem is that belief content is not identical to informational content. Belief content is individuated much more finely than informational content and the content of a belief can be false while information as defined by Dretske must always be true. Another problem is that not every information bearing state is a belief. A person's epidermal cells may carry information about the air temperature even though the person has no beliefs about the air temperature. Dretske handles these problems as follows:

1. The degree of intentionality problem: The first problem that Dretske faces in constructing a characterization of belief content in terms of informational content is that the two differ in what he calls their "order of intentionality". A kind of contentful state *S* exhibits *first order* of intentionality if it is possible that (a) All *F*'s are *G*, (b) an instance *s* of *S* has the content that *t* is *F* and (c) it is not the case that *s* has the content that *t* is *G*. A kind of state exhibits *third order* intentionality if the above holds where (a) is replaced by (a'); it is either nomically or analytically necessary that all *F*'s are *G*. It is clear that

informational states exhibit first but not third order of intentionality while belief states exhibit third order of intentionality.⁶ If neural state r 's being N carries the information that t is F , it will also carry the information that q where q is any proposition nomically or analytically implied by the proposition that t is F .⁷ r 's being N may also carry information that nomically implies that t is F . Of course, one can believe that t is F without believing everything implied by t is F or believing any proposition which implies that t is F . Also, if t is F and t is F^* are nomologically or analytically equivalent then if r carries the information that t is F it also carries the information that t is F^* , but it is possible to believe that t is F without believing that t is F^* . Informational content is quite coarse-grained. Belief content is clearly much more fine-grained.

Dretske deals with the first part of this problem by distinguishing the *semantic* content of a state from the rest of its informational content.

Structure S has the fact that t is F as its semantic content = (a) S carries the information that t is F and (b) S carries no other piece of information r is G , which is such that the information that t is F is nested (nomically or analytically) in $G(r)$. (p. 177)

The information that t is F is nomically (analytically) nested in the information that r is G iff r is G nomically (analytically) implies that t is F . It follows that the semantic content of $G(r)$ is the most *specific* information which it carries.

Dretske's initial suggestion is that belief contents are semantic contents of certain neural structures. By identifying belief content with the semantic content of a neural structure, it is possible to allow for a person's believing, for example, that t is a whale without believing that t is a mammal since his brain may contain a neural structure with the semantic content that t is a whale but contain none with the semantic content that t is a mammal.

Even if the preceding proposal is successful (I will later argue that it is not) there remains the problem that if $G(r)$ carries the information that p and, if q is nomically or analytically equivalent to p , then $G(r)$ also carries the information that q . It follows that it also carries the information that $p \& q$ and so cannot have either p or q as its semantic content. If $G(r)$ has a semantic content at all, it must be the conjunction of all the propositions which express the information that $G(r)$ carries.

Dretske's proposal for handling this problem can be illustrated with a simple example. An information system I contains two structures r and s . r can occupy a state G and the semantic content of r 's being G is that t is square. s is composed of two substructures, s_1 and s_2 . When the first occupies state H , its semantic content is that t is quadrilateral. When the second occupies state J , its semantic content is that t is equilateral. When s 's two components occupy H and J , s has the semantic content that t is quadrilateral and equilateral. It is possible for an organism's brain to contain both structures r and s and for the first to occupy G while the latter's components fail to occupy H and J and vice versa. According to Dretske, in the first case the organism believes that t is square but does not believe that t is an equilateral quadrilateral, even though the semantic contents of the two structures are identical.

Dretske's proposal seems to be this: If $G(r)$ is a belief state and if r is a *simple* structure then the belief content of r 's being G is its semantic content. If r is a complex structure then the belief content of $G(r)$ is some construct out of the semantic contents of the simple structures out of which r is composed. Two structures could then have the same semantic contents and yet correspond to different beliefs if the ways in which they are constructed from simpler structures differ.

Dretske doesn't develop the above account. One would like to know exactly what constitutes a simple structure, how complex structures can be built up from simpler ones, and how the content of a complex structure is determined by its composition. That is, one would like a specification of the syntax and semantics of cognitive structures.

2. The problem of misrepresentation: Another reason that belief content cannot simply be identified with informational content is that beliefs can be false while information content must be true. Here is how Dretske proposes to allow for misrepresentation:

Suppose that during the period L the system is exposed to a variety of signals, some of which contain the information that certain things are F , others of which contain the information that other things are not F . The system is capable of picking up and coding this information in analog form but, at the onset of L , is incapable of digitalizing this information. Suppose, furthermore, that during L the system develops a way of digitalizing the information that something is F . This semantic structure develops during L in response to the array of information bearing signals. Once this structure is developed, it acquires a life of its own, so to speak, and is capable of conferring on its subsequent tokens (particular instances of that structure type) its semantic content (the

content acquired during *L*) whether or not these subsequent tokens actually have this as their informational content. (p. 193)

Suppose, for example, that the pupil is being taught the concept *red*. During *the learning period* the pupil is exposed to signals which carry the information that *x* is red and to signals which carry the information that *x* is not red for various objects *x*. Dretske describes the situation during the learning period as one that is optimal for learning the concept. In this example presumably this means that the objects are fully in view, the lighting is normal and so on. By the end of the learning period the pupil may have developed a semantic structure, $G(r)$, which reliably has the semantic content that *x* is red. That is, by the end of the learning period all the tokens of $G(r)$ that occur under the conditions that prevail during the learning period have the semantic content that *x* is red. However, when the learning period is over, something may trigger *r* to be *G* even though *x* is not red. This is a case of misrepresentation. This particular token of $G(r)$ of course does not carry the information that *x* is red (since *x* is not red), but its *belief content* is that *x* is red. It inherits this content from the tokens which occurred during the learning period.

3. Not every structure with a semantic content is a belief. Dretske characterizes beliefs as "semantic structures that occupy an executive office in a system's functional organization" (p. 198). He does not fill in the details of this functionalist account of belief state. To do so would be to say precisely what distinguishes the role of belief states in the causation of behavior from the roles of other neural states. Dretske does suggest an interesting connection between the information-theoretic account of content and the role of belief in bringing about action. If Arabella waves her hand because she believes that Barbarella is across the street, her action is determined in part by the fact that her belief has the particular content that Barbarella is across the street. Dretske explains that for a structure *r* to cause *b* in virtue of the information it carries is for its possessing the property which carries that information to cause *b*. Thus, Arabella's belief that Barbarella is across the street causes her to wave in virtue of its content if $G(r)$ causes her waving and $G(r)$ has the semantic content that Barbarella is across the street.⁸

It will be useful to summarize our exposition of Dretske's account of belief.

- (D) *A* believes that *x* is *F* if there is a neural state $G(r)$ of *A*'s such that (a) $G(r)$ occupies the appropriate executive office in *A*'s functional organization and (b) By the end of the learning period for $G(r)$, $G(r)$ has the semantic content that *x* is *F*.⁹

Dretske observes that his account provides at least the beginning of a semantics for the language of thought. A neural structure *r* which is developed to carry the information that *t* is *F* for various objects *t* can be thought of as expressing the concept "is *F*". A complex neural structure *s* which is composed of substructures s_1 and s_2 might express the conjunction of the concepts expressed by s_1 and s_2 . While he offers no account of the semantics of quantifiers, names, or even relation expressions, the importance of this achievement, if in fact it is one, cannot be overestimated. It would show how the most puzzling feature of mental states, their representational powers, can arise out of the information processing capacities of neural structures.

DRETSKE'S ACCOUNT IN KFI DOESN'T WORK

Dretske's construction of intentional states from information carrying states is undoubtably attractive. But it just doesn't work. There are two difficulties that I will focus on. The first is that the semantic contents of neural states are not the contents of beliefs. They do not, contra Dretske, possess the requisite order of intentionality and do not correspond to the beliefs that we attribute to an organism to explain its behavior. The second is that the account of false belief fails to show how misrepresentation is possible.

Dretske's proposal is to identify, e.g., a frog's belief that there is a fly nearby with a neural structure in the frog. The belief has the content that *x* is a nearby fly because the semantic content that *r*'s being *G* acquired during the learning period is that *x* is a nearby fly. But it is implausible that any neural state, in frogs or in humans, has the semantic content that *x* is a nearby fly. The reason is that $G(r)$ will invariably carry information not only about distal stimuli, e.g., the fly, but also about proximal stimuli and other neural states which cause *r* to be *G*. $G(r)$ will also usually carry information about other events which are nomically connected to *x*'s being a nearby fly. $G(r)$ is some internal neural state and will undoubtedly have other neural states as

causal antecedents. It will carry information about these states in addition to carrying the information that there is a nearby fly. It follows that the semantic content of $G(r)$ will include information about neural states and so cannot be the intentional content that there is a nearby fly. Also, it will not be unusual for the presence of a nearby fly to carry information about other events. For example, there may be a law that whenever flies are present the air temperature is above freezing. It follows that $G(r)$ will carry the information that the air temperature is above freezing as well as the information that a nearby fly, so its semantic content cannot be that there is a nearby fly.¹⁰ The problem is that the semantic content of $G(r)$ will inevitably be much richer than the content of the belief with which r is identified. We might put it this way: the semantic content of r 's being G is the conjunction of all the information it carries, whatever states of the world it is correlated with, whether this information is employed by the frog or not. In contrast, belief content includes only that information which is information "for the frog". The problem is to show how one can extract this intentional content from the information in which it is embedded.

Dretske does address the problem that a neural state will carry information about proximal stimuli. He says:

A semantic structure's insensitivity to its particular causal origin, its muteness about the particular manner in which the information (constituting its semantic content) arrived, is merely a reflection of an important fact about beliefs. Our beliefs do not themselves testify to their causal origins. The fact that someone believes that Elmer died tells us nothing about how he came to believe this, what caused him to believe it. (p. 188)

Dretske is certainly correct in saying that the fact that someone is in a certain belief state may carry no information about how he came to be in that belief state. But it is just for this reason that his belief that x is F cannot be identified with $G(r)$ since this latter will inevitably carry information about matters other than x 's being F .

I suspect that the reason that Dretske does not see how difficult this problem is for his account is that he is not careful in specifying the feature of a neural state which carries the information used in constructing belief content. A neural structure r 's being G will have one semantic content; its being G^* will usually have a different semantic content. If we restrict ourselves to neurophysiological properties G , then $G(r)$ will carry information about other neurophysiological

events since there are *lawful connections* between events neurophysiologically described. Of course, there may also be features G^* of r such that $G(r)^*$ does not carry information about other neural structures or about proximal stimuli. For example, r 's being the belief that Elmer is dead may not carry such information since, as Dretske observes, beliefs do not testify as to their causal origins.¹¹ But even if r 's being the belief that Elmer died has the semantic content that Elmer died, Dretske cannot avail himself of this since he cannot employ intentional properties (e.g., the property of being a certain belief) in his account since he wants to construct a *physicalistic* reduction of intentionality.

Dretske does have a response to my criticism. Consider some neural structure r such that $G(r)$ is supposed to have the semantic content that Elmer is dead. It may be that r could come to be G via many causal routes; one route involving visual perception, another involving auditory perception, and so on. Each of these routes will involve different neural structures. There will be no telling from the fact that r is G that it took one route or another. So r 's being G in fact does not carry information about other neural structures.

There is a simple reply to this defense. Even if there are a large number of causal routes via which r might have come to have G it still is the case that one or another of those routes must have been followed. So $G(r)$ will include the information that route 1 or 2 or ... was followed. Furthermore, the belief content of the state type r 's being G is the semantic content it acquired during the learning period. At that time it is likely that there are a limited number of causal routes to $G(r)$. So its semantic content will include information that one of these limited number of routes have been taken. Whatever semantic content $G(r)$ might have, it will not be that *Elmer is dead* since it will include the information - that one or another of these routes was followed.

Perhaps there is some way to get around this objection at least for perceptual belief. Suppose that $G(r)$ is a perceptual belief with the perceptual object t , say a fly. We might identify the intentional content of $G(r)$, not with its semantic content but with the most specific information that it carries about the fly. The claim is that the most specific information that $G(r)$ carries about the fly does not include any information about other neural states, proximal stimuli or other events nomically connected to the fly; e.g. the air temperature

being above freezing. But this suggestion works only if the perceptual object of $G(r)$ can be nonintentionally specified and if the notion of information about t can be characterized. Dretske, in fact, does sketch an account of the objects of perception (p. 157). His view is that the object of a perceptual state is the event (or events) in the causal sequence leading to the production of the perceptual state which is the most proximal event about which the perceptual state carries information. For example consider the following causal sequence: a button is pushed, a bell rings, the air vibrates, A's ear drum vibrates, and perceptual state r occurs in A. According to Dretske r may carry information about the bell ringing but not about the particular vibrations in the air or the ear drum which caused r . The reason is that there are many possible vibrations varying in frequency and intensity which could have caused r . It is claimed that since r does not record the frequency and intensity of the vibrations these proximal stimuli are not the objects of the perception. Although r may carry the information that the button was pushed, the most proximal cause about which the r carries information is the bell so it is the perceptual object.

If Dretske's account of the object of perception was correct then we would have a reply to my objection that the informational content of a neural state is much richer than the content of any belief state. We could exclude information concerning neural states, proximal stimuli, and so forth from the content of $G(r)$ since none of these is the object of the state. However, it is clear that this account will not work for pretty much the same reason that belief contents are not semantic contents. r certainly will carry some information concerning the vibrations in the air and in the ear drum. It will at least carry information that the ear drum is vibrating since r does not occur (at least under normal conditions) unless A's ear drum is vibrating. It follows that the bell will not be the object of the perceptual state after all. The attempt to save the identification of belief content with semantic content by identifying the former with the most specific information concerning *the object of the belief* does not work.

I have been arguing that Dretske does not provide a satisfactory account of how belief content can be extracted from the informational content of neural states. Even if this problem could be solved there remains the problem of misrepresentation. Recall that Dretske's solution to this problem is to identify the belief content of a simple

cognitive structure $G(r)$ with the semantic content, say that x is F , that it has acquired by the end of the learning period. When r is G subsequent to the learning period it needn't carry the information that x is F and so might misrepresent that x is F .

If this account has any chance of succeeding Dretske must provide an explanation of what counts as the "learning period" for a cognitive structure. How can we tell when r is in the process of being developed and when its development is completed? He also must explain how it is possible for some tokens of a cognitive structure to carry the information that x is F while other tokens fail to carry this information.

What counts as the learning period for r is crucial in determining the belief content of $G(r)$. For example, suppose that during the period L the pupil is exposed to black cats. By the end of L he may have developed a structure r such that $G(r)$ has the semantic content that x is a black cat. (I am ignoring for now our arguments which showed that it is implausible that a neural state has these kinds of semantic contents.) Subsequently a grey cat triggers r to be G . This is a case of misrepresentation. But suppose that this last token is included in the learning period L' . Then $G(r)$ will have the semantic content that x is a black or grey cat and so will not misrepresent. Also, without a specification of the learning period we have no way of distinguishing misrepresentation from change of meaning. Suppose that during a learning period r acquires the content that x is a black cat. Some later tokens have the semantic content that x is a red cat. Are these misrepresentations? Perhaps the learning period has begun anew and r has acquired a new content.

Dretske simply gives no account of what constitutes the learning period. Of course most concepts are learned without the benefit of a teacher so we cannot identify the learning period as the period of instruction. Making matters more difficult is the fact that different cognitive structures presumably have different learning periods since there is no one period of time during which all concepts are learned. The most natural way of specifying the learning period for r is as the time it takes for the subject to be trained so that if r is G then x is F . But this specification assumes that r is G represents that x is F . It would be question begging for Dretske to appeal to such an account.

Dretske must provide a specification of the learning period which

yields an account of belief that meets our two adequacy conditions. It must be nonintentional and it must result in correct belief attributions. But he has given us no reason to think that this can be done.

There is another way of seeing the problem. Dretske assumes that it is possible for $G(r)$ to carry the information that x is F during the learning period even though some tokens of $G(r)$ that occur subsequent to the learning period fail to carry this information. How can this be if there is a law to the effect that when r is G then x must be F ?

I see one answer that Dretske might give to this question. Recall that the information which an event carries is relative to channel conditions. During the learning period conditions are supposed to be optimal for learning the concept. As Dretske says "not only is one exposed to F 's and non- F 's but the fact that F 's are F and non- F 's are not F is made perceptually *obvious* or *evident*" (p. 194). For a concept like x is a *black cat*, presumably this means that the pupil is exposed to a variety of black cats under good lighting, for a sufficient length of time, etc. The suggestion is that we identify the channel conditions for a concept with the optimal conditions for learning it. Suppose that under these conditions the pupil acquires a structure r such that $G(r)$ has the semantic content that x is a black cat. Subsequently, the pupil catches a glimpse of an animal as it dashes into the bushes. It is a large rat that triggers r to be G . Since the channel conditions that obtained during the learning period no longer obtain, conditions are no longer optimal, it is possible for r 's being G to occur without carrying the information that x is a black cat. (It might carry the information that x is a large rat or it might not.)

For this account to work we need a way of distinguishing the "optimal" conditions that hold during the learning period from the conditions that may hold subsequent to it. But the optimal conditions for learning one concept may differ from those for learning another. What counts as optimal depends on the concepts *content*. Conditions which make it "obvious or evident" that x is a black cat differ from those which make it evident that t is a star. We cannot tell whether conditions are optimal ones for r until we know what r is supposed to mean. Thus our characterization of the channel conditions relative to which we determine the information carried by a concept by the end of the learning period relies on semantic notions.

DRETSKE'S NEW ACCOUNT OF MISREPRESENTATION

My criticisms show that Dretske's information-theoretic account of belief content fails both conditions of adequacy. The contents it assigns to belief states are not belief contents. The account of misrepresentation ultimately relies on a semantic characterization of a concepts learning period. Dretske himself seems to be dissatisfied with his account of misrepresentation. In a recent paper he suggests a different account of misrepresentation.¹² In this paper he doesn't attempt a reduction of belief content but is content to show how a state can be said to misrepresent without presupposing any intentional or semantic notions. The more recent suggestion is that what a brain state (or other internal state of an organism) represents is determined by the information which it is *designed* to carry. Dretske brings in teleological considerations in order to help specify representational content. He defines the "functional meaning" of a structure *r*'s being in a state *G* as follows:

r's function is to indicate the conditions of *w* and the way it performs this function is, in part, by indicating that *w* is *F* by it (*r*'s) being *G*. (p. 7)

The idea is that $G(r)$ functionally means that *w* is *F* if when *r* is functioning properly $G(r)$ carries the information that *w* is *F*. Misrepresentation can occur when *r* is not functioning properly.

According to the account in KFI the content of $G(r)$ is determined by the correlations which obtain during the learning period. On the new account it is the correlations that obtain when *r* is functioning properly that determine the content of $G(r)$. This seems to be an improvement since, as I argued previously, the distinction between the period during which a concept is learned and when it is used is question begging. But, as Dretske recognizes, his new account requires that there be a nonintentional characterization of *r*'s information carrying function. One might worry that we see a structure as having the function to carry certain information not because this is an objective feature of the structure but because of the use that we make of the structure. For example, if the column of mercury registers 70 degrees when it is 80 degrees because the mercury is stuck we will say that the thermometer misrepresents. But certainly this depends on the fact that we use the thermometer to measure temperature. What counts as proper functioning is relative to our purposes. Dretske

concludes that the ability of the thermometer to misrepresent (and so to represent) is dependent on human beliefs and purposes. It is not "original representation".

Dretske observes that the natural place to look for original representation is in organisms which contain structures that appear to have evolved, because they play a vital information gathering role in the species' adaptation to its environment. He considers the example of certain anaerobic bacteria which contain internal magnets called "magnetosomes". The magnetosomes of those bacteria that live in the northern hemisphere point to magnetic north and cause the bacteria to swim away from the surface. The survival value of the structure appears to be that it helps the bacteria avoid oxygen rich, hence toxic, surface water. Dretske considers the claim that the function of the magnetosome is to carry information about the direction of oxygen rich water. Supporting this is the fact that the bacteria need to avoid oxygen rich water to survive and it is plausible that they would not have evolved to contain magnetosomes unless these function to detect the direction of oxygen rich water. If the south pole of a bar magnet is placed near one of these bacteria and it moves upward (toward local magnetic north and in the direction of oxygen rich water) it is *misrepresenting* the direction of oxygen rich water. At first, this looks like a case of natural misrepresentation. But Dretske realizes that the information that a structure normally carries and the needs which are satisfied by the information do not uniquely determine the structure's information carrying-function (if it has any). In the case under discussion, another account is that the function of the magnetosome is to carry information about the direction of local magnetic north. It so happens that in its usual environment the direction of local magnetic north is correlated with the direction of oxygen rich water. If this is the magnetosome's function then a bacterium allegedly fooled by the bar magnet really does not misrepresent the direction of oxygen rich water since it is not the function of the magnetosome to carry information about the direction of oxygen rich water but about the direction of local magnetic north. It is carrying out this function perfectly. As Dretske says, we "need some principled way of saying what the natural function of a mechanism is".

Dretske next considers a (hypothetical) organism, *O*, which contains a structure *r* such that $G(r)$ carries the information that *x* is *F* but which is unlike the bacteria in that there are two ways in which *O* can

detect F s. For example, consider an animal capable of detecting the presence of a lion either by sight or sound. Figure 1 illustrates the situation.

F is the property of being a lion, s_1 and s_2 are respectively the proximal stimuli, appearing to have a large mane and a roaring sound; and r is a cognitive structure. When r is functioning properly it is caused to be in state G either by s_1 or by s_2 . When conditions are normal only lions appear to have large manes and only lions roar. So if r is G then there must be a lion present. When circumstances are not normal, misrepresentation may occur. (e.g., someone dresses up in a lion suit and roars, causing O mistakenly to believe that a lion is present.) In the case of the bacterium there is not genuine misrepresentation since the magnetosome represents the direction of local magnetic north, not the direction of oxygen rich water. Consider the analogous objection that in the present case when O is confronted with someone dressed in his lion suit, $G(r)$ doesn't misrepresent that a lion is present but correctly represents the occurrence of one of the proximal stimuli, e.g., a roaring sound. Dretske at first suggests that this objection is unfounded. He points out that $G(r)$ does not represent the occurrence of a roaring sound since it does not carry this information even under normal conditions. That is, when r is func-

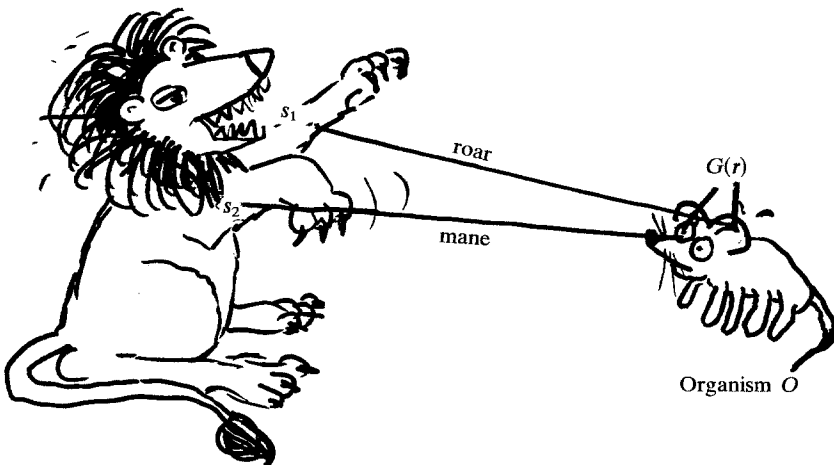


Fig. 1.

tioning normally, it is possible for it to be G even though no roaring sound occurs. Similarly, r 's being G doesn't represent any other proximal stimuli that normally occasions it. So it appears that Dretske has discovered a principled way of eliminating the proximal conditions as what is represented and so rescues his account of misrepresentation.

However, there is an obvious objection, analogous to the objection we made against the identification of belief content with semantic content, which Dretske himself raises against this proposal. It is that $G(r)$ instead of representing that a lion is present, represents the disjunctive proposition that either s_1 or s_2 has occurred. If this were so then we no longer would have a genuine case of misrepresentation since $G(r)$ *correctly* represents that s_1 or s_2 has occurred. As long as this alternative cannot be dismissed, we do not have an unequivocal case of misrepresentation.

Dretske's response to this is to consider an organism capable of associative learning. Suppose that originally $G(r)$ can be caused either by s_1 or by s_2 , as previously. But the organism is capable of "learning" to associate a new proximal stimulus, s_3 , with $G(r)$. Dretske further supposes that

.. it becomes clear that there is virtually no limit to the kind of stimulus that can acquire this "displaced" effectiveness in triggering r and subsequent avoidance behavior.

He concludes that in this kind of situation it cannot be said that r 's being G represents any of the disjunctions $s_1 \vee s_2 \vee \dots$, since there is no end to the disjunctions which can come to be associated with r 's being G . So, none of the disjunctions can be what $G(r)$ represents.

We have been assuming that although various s_i can come to trigger $G(r)$ those which do are all correlated, *under normal conditions*, with the presence of a lion. Dretske doesn't tell us exactly what constitutes normal conditions but presumably he would count the presence of holographic images of lions or tape recorded roars as not normal. He claims that $G(r)$ represents the presence of a lion since this is all that is in common among the s_i that come to be correlated with $G(r)$. In "abnormal" circumstances, a stimulus s_k can occur and cause r to be G even though no lion is present. According to Dretske, this is a genuine case of misrepresentation. One cannot now object that $G(r)$ actually represents some disjunction of proximal stimuli $s_1 \vee s_2 \vee \dots$ since even in normal circumstances it doesn't carry the information that $s_1 \vee s_2 \vee \dots$.

Has Dretske actually constructed a naturalistic account of misrepresentation? He has told us that $G(r)$ represents for O that x is F if it is the function of $G(r)$ to carry information that x is F . He doesn't provide a positive characterization of "the function of $G(r)$ is to carry the information that p ". But he does give us a way of discovering what information it is *not* the function of $G(r)$ to carry. In our example, it was not the function of $G(r)$ to carry the information that s_1 or s_2 occurred since O could learn to associate other stimuli with $G(r)$. To help evaluate his proposal let's examine the following positive account of representation which is based on his discussion:

- (DD) $G(r)$ represents that x is F iff the most specific information common to every D -possible token of $G(r)$ which occurs when conditions are normal is that x is F and O needs the information that x is F .¹³

It is clear that in determining what $G(r)$ represents, we must consider not only those tokens of G which occur during O 's life, since these will be associated with a finite, perhaps small, set of stimuli and we would not have eliminated the claim that stimuli $G(r)$ represents the disjunctive proposition that one of these stimuli occurred. Dretske's suggestion for eliminating this alternative is that we consider certain counterfactual tokens of r 's being G . These are the D -possible tokens of $G(r)$. These counterfactual tokens are caused by various stimuli other than those which cause actually occurring tokens. By including these tokens in the determination of the informational function of $G(r)$ one might hope that the representational content of $G(r)$ will not include information about proximal stimuli cause r to be G .

This new account of representation appears to be an improvement over the one found in KFI. By considering the D -possible tokens of r 's being G , the new account seems to avoid our objection to semantic content. We cannot object to the new account that $G(r)$ carries the information that a certain disjunction of stimuli or neural states occurred (making this part of its semantic content), since there may be D -possible tokens which are not caused by any of these stimuli or neural states. The new account does not appeal to the dubious distinction between the period during which a concept is learned and period when it is used. Instead, misrepresentation is accounted for by the failure of the structure to function normally or by conditions being abnormal. Since my two major objections to the old account concer-

ned problems with the learning period and semantic content it will be interesting to see if the new account avoids similar objections.

If (DD) is to be a successful naturalistic account of representation, then it must be possible to characterize *O*'s informational needs, normal conditions and functioning, and *D*-possible tokens of *G(r)*, in nonintentional terms. I have some worries about the first two of these but the main difficulty is in characterizing the *D*-possible tokens of *G(r)*. As we observed, we do not want to restrict the *D*-possible tokens to those which are caused by whatever stimuli *O* actually learns to associate with *r*'s being *G*, since then the representational content of *G(r)* would include that one of these stimuli occurred. So we include possible tokens of *G(r)* which are caused by other stimuli. But given the plasticity of neural structures, it is certainly possible for *r*'s being *G* to be caused by stimuli which are not themselves normally associated with *x*'s being *F*. That is, there are counterfactual situations in which something other than *x*'s being *F* causes a stimuli which causes *r* to be *G*. If we count such tokens as *D*-possible then (DD) will not attribute to *G(r)* the content that *x* is *F*. The problem is to distinguish those tokens which, while associated with various stimuli, still represent that *x* is *F* from those that represent some other concept, and to do this in a way that does not presuppose semantic or intentional notions. We cannot, for example, say that the *D*-possible tokens of *G(r)* are the ones which represent that *x* is *F* (or are caused by *x*'s being *F*) since this obviously appeals to what *G(r)* represents. Dretske offers no hint as to how to characterize the set of *D*-possible tokens without employing semantic notions and I see no way in which it can be done. In any case, until he tells us precisely which tokens of *G(r)* can be considered in determining what *G(r)* represents he has not produced an account of misrepresentation.

FODOR'S PSYCHOSEMANTICS

Jerry Fodor has constructed a sophisticated version of the informational theory of representation which, although inspired by Dretske's approach in KFI, does not at first seem to succumb to the same objections.¹⁴ According to Fodor, to believe that *p* is to bear a certain computational relation *B** to a mental representation *S* which means that *p*. Mental representations are sentences in a language of thought which have a syntax and a semantics. Fodor supposes that the

relation B^* can be characterized syntactically or at least without recourse to intentional or semantic notions. To make matters more graphic, he adopts a suggestion of Schiffer's and imagines that to bear B^* to M is to place a token of M in a YES box. It is part of the function of the various cognitive systems – memory, perception, reasoning, etc. – to put tokens of representations in and take them out of the YES box. Even if we possessed an account of the internal workings of the YES box we would lack semantics for the language of thought. The task of what Fodor calls “psychosemantics” is to characterize the meanings of mental and thereby the contents of belief.

Fodor holds that the meaning of a mental representation is a product of two factors: its truth conditions and its inferential role. He does not provide much illumination concerning inferential role. As he says, in appealing to it he is mostly whistling in the dark. The important feature of inferential role for our discussion is that it be characterized without appeal to semantic or intentional concepts. The following might do: S 's inferential role is the effect that S 's being placed in the YES box would have on other representations being in the YES box and the effect that other representations being in the YES box would have on S 's being in the YES box.¹⁵

Fodor holds a two factor theory because he thinks that two mental representations S and S' may have the same truth conditions and yet aB^*S and aB^*S' may attribute beliefs with different contents. For example, according to Fodor, “The Morning Star is bright” and “The Evening Star is bright” (imagining that these are mental representations) have identical truth conditions but may differ in their inferential roles.

Fodor sees the main problem of psychosemantics as constructing a naturalistic account of truth conditions. He begins by defining “the entry condition for a mental representation M ” as

... that state of affairs such that under conditions of normal functioning the organism's cognitive system puts M in the YES box iff the state of affairs obtains. (p. 37)

The entry conditions for M are the conditions that are causally necessary and sufficient for M 's being in the YES box when the organism's cognitive systems are functioning normally. Fodor first proposes that M 's truth conditions are its entry conditions:

... the entry condition function for a system of mental representations is coextensive

with its truth definition. That is, for each mental representation M , S is the truth condition for M iff S is the entry condition for M . Or, rather, something stronger: what makes S the truth condition M — the precise fact in virtue of which S counts as the truth condition for M — is that S is the entry condition for M . (p. 44)

Fodor's account of belief content can be formulated as follows:

- (F) A believes that p iff there is a representation S such that a) $aB * S$ and b) S has inferential role appropriate to " p " and c) when A 's cognitive systems are functioning normally p is causally sufficient and necessary for $aB * S$.

Fodor's account is similar to Dretske's in that a belief has the content that p only if the belief carries the information that p under certain "optimal" conditions. But there are important differences between the accounts. One difference is that in addition to requiring that S carry the information that p under optimal conditions, Fodor requires that under optimal conditions p is causally sufficient for $aB * S$. In one respect this is an improvement over Dretske since, as we saw, when r 's being G carries the information that a fly is nearby it will also carry the information that the air temperature is above freezing. The consequence was that the content of $G(r)$ could not be precisely that a fly is nearby. Fodor doesn't have this problem since the air temperature's being above freezing might not be causally sufficient for $aB * S$ even if it is necessary. However, he still has the problem that representations which are analytically or nomically equivalent will have the same truth conditions. Presumably he hopes to deal with this by appealing to differences in inferential role. However, as I have argued elsewhere, no proponent of two-factor theories has shown how truth conditions and inferential role can be combined into meaning.¹⁶

A second difference is the way optimal conditions are characterized. In Dretske's KFI account optimal conditions (the C_n in $[IS]$) are the channel conditions that hold during the learning period. Fodor characterizes optimal conditions in terms of the proper functioning of the cognitive systems. In this it bears a similarity to Dretske's revised account.

Fodor's claim that it is the function of the cognitive systems to put S in the YES box iff its truth condition is satisfied is supposed to be an empirical hypothesis supported by evolutionary considerations.¹⁷ It is intended to be akin to "it is the function of the heart to pump blood". At first glance it is incredible that it is a function of my cognitive

system to place the representation "Jerry Fodor is in Palo Alto" in my YES box iff Jerry Fodor is in Palo Alto. Most of Fodor's paper is devoted to dispelling this initial incredulity. But I will argue that (1) there are reasons to doubt that it is the function of the cognitive systems to result in true beliefs; (2) in any case when functioning properly is understood biologically the beliefs (*F*) attributes to a person are certainly not the beliefs that any person has; and (3) when one tries, as Fodor does, to improve on (*F*) so that it yields more adequate belief attributions the improvements rely on intentional and semantic concepts.

(1) There are some considerations which mitigate against the general claim that it is the function of the cognitive systems to yield *true* beliefs. When apparently functioning normally human perception and memory are subject to errors, as various well known examples of perceptual illusions found in psychology texts show. Recent research in human reasoning summarized by Nisbett and Ross seems to reveal that it too is prone to "illusions".¹⁸ These authors suggest that normal human reasoning involves the use of heuristics which, although they generally yield correct or nearly correct beliefs in everyday contexts, can result in incorrect beliefs outside of such contexts. One can easily imagine natural selection resulting in cognitive systems that efficiently acquire mostly true beliefs in one area at the expense of error in other areas.¹⁹ Fodor observes that many animals apparently "jump to conclusions" when it comes to the identification of predators. A bird hunted by hawks attempts to escape at the mere flutter of a black cloth. It is natural to describe the bird as believing that a hawk is nearby and to say that its cognitive system is so designed to reach this conclusion even on the basis of such slight evidence. Part of Fodor's response is to redescribe the case as one in which the bird assigns a low probability to the presence of a hawk but because of the extremely low utility of being captured by a hawk, the bird attempts to flee. But this is not very convincing since it is implausible that humans, let alone birds, act to maximize expected utility.

(2) Fodor's theory of belief has an incredible consequence. As he himself observes, it follows from (*F*) that "under normal conditions, everybody is omniscient" (p. 48). That is, when an organism's cognitive system is functioning properly it will believe that there is a fly present iff a fly is present and that Jerry Fodor is in Palo Alto iff Jerry Fodor is in Palo Alto and so on. Fodor remarks that the omniscience is

limited to beliefs that the organism is capable of entertaining, but this is not much consolation. If “the normal functioning of the cognitive systems” is understood in terms of normal biological and psychological functioning then it is simply not true that a person whose cognitive systems are functioning normally will believe that p iff p . Fodor doesn’t tell us what it is for a human cognitive system to be functioning properly. But let us suppose that Arabella’s eyesight is 20–20, her hearing perfect, her memory excellent, her college board scores all 800 and so on. Jerry Fodor is in Palo Alto but despite the excellent condition of her cognitive system Arabella doesn’t believe it. It follows from (F) that Arabella has no mental representation in her language of thought with the truth condition that Jerry Fodor is in Palo Alto. So Arabella does not and cannot believe that Jerry Fodor is in Palo Alto. The conclusion is the same for most other beliefs. This is intolerable.

(3) Fodor is aware of the preceding difficulty and modifies his account to deal with it. He suggests that the reason that Arabella fails to believe that Jerry Fodor is in Palo Alto even though her cognitive systems are all functioning is that certain “epistemic appropriateness” conditions for that belief are not satisfied. This leads to the following modification of (F):

(F*) A believes that p iff there is a representation S such that a) $aB * S$ and b) S has an appropriate inferential role and c) when A ’s cognitive systems are functioning normally and *epistemic appropriateness conditions for S hold*, p is causally sufficient and necessary for $aB * S$.

What are epistemic appropriateness conditions? Fodor gives this explanation:

What I am committed to is just that we are omniscient in the circumstances that are relevant to the normal functioning of our cognitive apparatus, where this includes the satisfaction of all pertinent conditions of epistemic appropriateness. To put it vulgarly, I am committed to: rub our noses in the fact that P and (if we can frame the thought that P) we’ll come to believe that P . (p. 50)

He suggests that the bird of a few paragraphs back may have believed that a hawk was nearby after all, since the epistemic appropriateness conditions for that belief were not satisfied. Perhaps if the bird had inspected the cloth it would not have fled. The problem was its nose

was not rubbed in the fact that it was a cloth not a hawk fluttering above it.

The modification is a major one. While it is plausible that the functions of the cognitive system can be characterized nonintentionally it is quite the opposite with regard to epistemic appropriateness conditions. It is difficult to see how the epistemic appropriateness conditions of belief states can be characterized without appeal to intentional and/or semantic concepts. Notice that any case of false belief must be explained on Fodor's account either by failure in the cognitive systems or by the nonsatisfaction of epistemic appropriateness conditions. Even though her cognitive systems are functioning properly, Arabella fails to believe that Venus is above the horizon when it is. The reason is that it is daytime and Venus is not visible. It seems to follow that its not being daytime is an epistemic appropriateness condition for "Venus is above the horizon." Arabella fails to believe that the animal in front of her is a bear although it is. The reason is that it is dusk and Arabella takes the animal to be a dog. Apparently among the epistemic appropriateness conditions for "The animal in front of me is a bear" is that it is not dusk. There obviously is no one epistemic appropriateness condition for all beliefs. Instead, epistemic appropriateness conditions vary and the epistemic appropriateness conditions for one belief might exclude those for another. Fodor provides no account of how one is to determine what epistemic appropriateness conditions are to be associated with a representation. It is the case that if *a*'s cognitive system is functioning properly and *aB** "Jerry Fodor is in Palo Alto" when Jerry Fodor is not in Palo Alto (or Jerry Fodor is in Palo Alto although not *aB** "Jerry Fodor is in Palo Alto"), then *a*'s situation is not epistemically appropriate. But this doesn't isolate epistemic appropriateness conditions. And, more significantly, it presupposes that we already know the meaning of *a*'s mental representations. If we didn't we wouldn't know when *a*'s beliefs are false. Furthermore, our examples make it clear that the epistemic appropriateness conditions for *S* depend on what *S* is about; what it means. If this is correct, then Fodor's appeal to epistemic appropriateness conditions violates the second of the adequacy conditions on a naturalistic account of content which prohibits the employment of intentional and semantic concepts.

Fodor might reply to this that I have not shown that it is impossible to specify epistemic appropriateness conditions without presupposing

semantics. Perhaps there is some way of associating appropriateness conditions with mental representations based on their inferential roles or other syntactic features. (Though this task appears as Herculean as the original task of associating truth conditions with representations.) However, there are some considerations which strongly suggest that no specification of entry conditions can satisfy our two adequacy conditions.

First, I want to employ a famous example of Tyler Burge's.²⁰ Arabella has a pain in her thigh and says "I have arthritis in my thigh". Her doppelganger, twin-Arabella, also has a pain in her thigh and says "I have a pain in my thigh" but she correctly believes that she has arthritis. The relevant difference between the two is that Arabella speaks a language in which "arthritis" is confined to inflammations of the joints while twin-Arabella speaks a language in which "arthritis" refers to inflammations of the joints or thighs. Burge plausibly claims that Arabella's belief is false while her twin's belief is true. Since Arabella's belief is false, it follows from (F^*) that some epistemic appropriateness condition for "I have arthritis in my thigh" is not satisfied. But what could it be? Whatever it is it must involve the meaning of "arthritis" in her language. Since what she fails to know is the meaning of "arthritis" in her language, a likely candidate is that knowing the meaning of "arthritis" is part of the epistemic appropriateness conditions of "I have arthritis in my thigh". If this is correct then epistemic appropriateness conditions, at least for some beliefs, are irreducibly semantic.

The Burge example is intended to show that the epistemic appropriateness conditions for certain beliefs involve knowledge of the semantics of one's language. This obviously undermines Fodor's attempted reduction. Fodor could reply by questioning Burge's claim that Arabella's belief is false (perhaps it's not correct to say that she believes that she has arthritis in her thigh) or by restricting his account to beliefs involving concepts which are not susceptible to Burge-style arguments. But the example reveals another problem with (F^*) that is quite distinct from Burge's argument. It is that the epistemic appropriateness conditions for a particular belief involve having and not having certain other beliefs. This is made clear in the following two examples:

(a) Arabella is looking right at a photograph from SLAC which depicts beta decay. Yet she fails to believe that it depicts beta decay.

The explanation is not any failure in her cognitive systems or her not being in the right place but rather that she fails to know anything about linear accelerators, sub-atomic particles and their photographs. The example shows that among the epistemic appropriateness conditions for "this depicts beta decay" is having a great many other beliefs. Recall that Fodor characterizes the epistemic appropriateness conditions for a belief as having one's nose rubbed in the fact expressed by the belief. One's nose cannot be rubbed in the fact that this is a case of beta decay without rubbing one's nose in a great many other facts first.

(b) Arabella is looking straight ahead at Barbarella. Although her cognitive systems are functioning properly, etc., she does not believe that she is looking at Barbarella since she falsely believes that she is looking at a holographic image. This shows that not believing that one is looking at a holographic image is part of the epistemic appropriateness conditions for "I am looking at Barbarella". Of course, it is not part of the epistemic appropriateness conditions for "I am looking at a holographic image".

These two examples depend on the *fact* that belief formation is holistic. Whether or not one will put *S* into one's YES box given certain stimuli depends on what other representations are already there. Example (a) shows that the epistemic appropriateness conditions for some beliefs include having certain other beliefs. Example (b) shows that they exclude having certain other beliefs. Furthermore, whether having or not having certain representations in one's YES box is included in the epistemic appropriateness conditions for *S* depends on what those representations mean. The only hope I can see for Fodor's saving (F^*) in the face of this is to find some way of specifying these belief states *nonsemantically*. Good luck!

CONCLUSION

I have examined the most promising of the information-theoretic accounts of belief and argued that they are not successful. Although I cannot *prove* that no information-theoretic semantics is correct, the way in which Dretske's and Fodor's accounts fail suggest that none will work. There are really two problems. First, informational content is much more coarse-grained than belief content. Neither Dretske's atomism nor Fodor's appeal to conceptual role overcomes this prob-

lem. Second, in attempting to specify the conditions C_n under which belief content is identified with (or included in) informational content in a way that yields plausible belief attributions the advocates of (*IS*) are forced to invoke intentional and semantic notions. Most of the argument of this paper consisted in making this plain. The appeal to intentional and semantic notions in characterizing the C_n seems inevitable since, as we argued in discussing Fodor, the conditions under which a belief state carries information that corresponds to its content clearly depends on that content. None of the proponents of (*IS*) have found a way of specifying these conditions without in some way appealing to that content and I suspect there is none to be found.²¹

NOTES

¹ There is an important difference between standard versions of behaviorism and the identity theory. The former identifies mental states with behavior or dispositions to behave, and claims that the identification is conceptual. The latter identifies mental states with brain states and claims that the identification is contingent.

² This point has been made by many including John Searle (1980), Jerry Fodor (1981), and Hilary Putnam (1981).

³ See Fodor (1984a,b), Dretske (1981, 1986), and Stalnaker (1984). An earlier version of this kind of account is found in Stampe (1977). Barwise and Perry (1983) seem to presuppose an information-theoretic account.

⁴ Those who advocate the information-theoretic accounts of belief considered here would probably be satisfied with something less than a complete reduction that identifies the belief that p with a physically specifiable state that applies to all possible believers. An account that applies to humans, even just to some human beliefs, would be a sufficient achievement. This slight relaxing of the aims of the account does not affect the criticisms I will make of it.

⁵ For a discussion of the difficulty of interpreting "probability" in Dretske's definition of information see Loewer (1983).

⁶ Not everyone agrees that belief exhibits third order intentionality. Stalnaker (1984) apparently disagrees and then attempts to explain away appearances to the contrary. However Stalnaker does hold that belief exhibits second order intentionality — one can believe p without believing that q even when p and q are nomically equivalent — and this creates a problem for his and other information-theoretic accounts.

⁷ p nomically implies q iff $p \& CC \& laws$ imply q but $CC \& L$ do not imply q .

⁸ Dretske develops his account in (1986b) to explain how it is that a belief's being the belief that p can cause an action.

⁹ (*D*) is limited to de re belief attributions of the form: "A believes of x that it is F " where " F " is a simple concept. Dretske does not explain how we can tell whether or not a belief attribution in English attributes a belief involving a simple or a complex concept.

¹⁰ Dretske holds that if $G(r)$ carries the information that $F(s)$ and $F(s)$ carries the

information that $H(t)$ then $G(r)$ carries the information that $H(t)$. He calls this "the Xerox principle". It may be that he should not hold this principle since the channel conditions may change from $G(r)$ to $F(s)$. In any case, in this example the presence of a fly carries the information that the air temperature is above freezing and $G(r)$ carries the information that there is a fly so $G(r)$ carries the information that the air temperature is above freezing.

¹¹ This doesn't seem to be exactly right. A 's believing that it's raining downtown might in appropriate circumstances carry the information that she has been downtown. It always carries the information that A is alive.

¹² See Dretske (1985). Dretske doesn't make any reference to his account of misrepresentation in KFI in this paper.

¹³ Dretske does not give DD as an account of representation but it is implicit in his account of misrepresentation. It is impossible to evaluate an account of misrepresentation without an account of representation.

¹⁴ Fodor (1984b). Various versions of this paper have been in circulation (page references are to the latest version). Fodor has expressed doubts about the teleological claims of his account in Fodor (1985, 1986). In Fodor (1986) he develops yet another different account of representation and misrepresentation. While the specific objections I make in this paper do not apply to his new account I argue (Loewer, 1987) that they also are ultimately question begging.

¹⁵ The view that meaning can be analyzed in terms of two components, truth conditions and conceptual role, has become popular recently. See Field (1977), McGinn (1982), and Block (1985). For criticisms of this approach see Schiffer (1985) and LePore and Loewer (1985).

¹⁶ See LePore and Loewer (1985).

¹⁷ Fodor holds that teleological contexts are intensional but can be analyzed without appeal to any intentional or semantic notions. He speculates that some analyses in terms of counterfactuals concerning natural selection, etc. For incisive criticisms of Fodor's use of evolutionary theory to support his teleological assumptions see Silvers (1985).

¹⁸ See Nisbett and Ross (1980).

¹⁹ See Stich (1985) for persuasive arguments that natural selection does not always select true believers.

²⁰ See Burge (1979).

²¹ Paul Boghossian helped me get clear on how the holism of belief fixation undermines Fodor's account.

²² I would like to thank Davis Baird, Paul Boghossian, Mike Costa, Gary Gates, Ernest LePore, Brian McGlaughlin and Dion Scot-Kakures for helpful discussions of information-theoretic semantics.

REFERENCES

- Barwise, J. and J. Perry: 1983, *Situations and Attitudes*, Cambridge, Mass.
 Block, N.: 1985, 'Advertisement for Semantics for Psychology', *Midwest Studies in Philosophy* 10, 615-678.
 Burge, T.: 1979, 'Individualism and the Mental', *Midwest Studies in Philosophy* 4, 73-121.

- Dretske, F.: 1981, *Knowledge and the Flow of Information*, Cambridge, Mass.
- Dretske, F.: 1986a, 'Misrepresentation', to appear.
- Dretske, F.: 1986b, 'Explaining Behavior', to appear.
- Field, H.: 1977, 'Logic, Meaning, and Conceptual Role', *Journal of Philosophy* 74, 379-409.
- Fodor, J.: 1981, *Representations*, Cambridge, Mass.
- Fodor, J.: 1984a, 'Semantics, Wisconsin Style', *Synthese* 59, 1-20.
- Fodor, J.: 1984b, 'Psychosemantics', widely circulated ms.
- Fodor, J.: 1985, 'Causal Theories and the Problem of Error', ms.
- Fodor, J.: 1986, *Psychosemantics*, Cambridge, Mass.
- Loewer, B.: 1983, 'Information and Belief', *Behavioral and Brain Sciences* 5, 79-80.
- Loewer, B.: 1986, 'Why the Problem of Misrepresentation is So Hard', to appear. A version was read to the Society for Philosophy and Psychology, June 1986, Baltimore, Maryland.
- Loewer, B. and E. LePore: 1986, 'Dual Aspect Semantics', in LePore (ed.), *New Directions in Semantics*, London.
- McGinn, C.: 1982, 'The Structure of Content', in A. Woodfield (ed.), *Thought and Object*, Oxford.
- Nesbitt, J. and J. Ross: 1980, *Human Inference: Strategies and Shortcoming in Human Judgment*, Prentice-Hall, Englewood Cliffs.
- Putnam, H.: 1981, *Reason Truth and History*, Cambridge.
- Searle, J.: 1980, 'Minds, Brains, and Programs', *Behavioral and Brain Sciences* 2.
- Silvers, S.: 1985, 'Natural Teleology', Read to the Society for Philosophy and Psychology, June, 1985, London, Ontario.
- Stalnaker, R.: 1984, *Inquiry*, Cambridge, Mass.
- Stampe, D.: 1977, 'Toward a Casual Theory of Linguistic Representation', *Midwest Studies in Philosophy* 2, 42-63.

Department of Philosophy
University of South Carolina
Columbia, SC 29208
U.S.A.