The Role of Mental Meaning in Psychological Explanation[1]

Robert Cummins

University of Arizona

Abstract

In Explaining Behavior, Dretske sets out to explicate the causal role of content in psychological explanation. This project is motivated by two assumptions: (i) the role of semantic content in folk-psychological explanation just is (or is our best line on) the explanatory role of semantic content in scientific cognitive psychology, and (ii) semantic content cannot have a serious explanatory role to play in psychology unless it has a causal role to play. Both assumptions are challenged in this paper. The explanatory role of semantic content in most contemporary cognitive science is not causal and has little to do with the explanatory role of reasons in folk-psychology.

Introduction.

In his recent book Explaining Behavior,[1] Fred Dretske sets out an extended and admirably clear account of the role of mental meaning in the explanation of behavior. The account is unique in being the only well-informed current attempt I know of to take seriously and attempt to explicate the idea that the semantic content of a mental state is causally relevant to the explanation of behavior. Dretske's effort can (and should), therefore, be read as, among other things, a response to what I call Stich's Challenge:.2 Given that, by definition, the semantic properties of a representation in a computational system are not relevant to its causal role, how can a computationalist take semantic content to be relevant to psychological explanation?.3

Dretske's account is explicitly directed at ordinary commonsense explanation by reasons--i.e., at 'folk-psychological' explanation, as it has come to be called. Dretske obviously takes this sort of explanation very seriously. He is not the kind of philosopher who would write a book

[1] In B. McLaughlin, ed., *Dretske and his Critics.* Blackwell, 1991, pp. 102-118.

about a kind of explanation if he did not think that kind of explanation is on the right track. Dretske, along with Fodor and many others,.4 assumes that commonsense explanation by reasons is, if not yet science, at least near to it, and bound to be (perhaps in regimented form) at the core of any serious psychology of cognition. Dretske is thus (again, along with Fodor and many others) making an assumption I call the Intentionalist Assumption:

> (IA) the role of semantic content in folk-psychological explanation just is (or is our best line on) the explanatory role of semantic content in scientific cognitive psychology.

Dretske is also making what I will call the Causal Assumption:

> (CA) semantic content cannot have a serious explanatory role to play in psychology unless it has a causal role to play.

He writes:

> Something possessing content, or having meaning, can be a cause without its possessing that content or having that meaning being at all relevant to its causal powers. . . . If meaning, or something's having meaning, is to do the kind of work expected of it--if it is to help explain why we do what we do--it must, it seems, influence the operation of those electrical and chemical mechanisms that control muscles and glands. Just how is this supposed to work? This obviously, is as much a mystery as the interaction between mind stuff and matter.5

These assumptions--the ones I've labeled (IA) and (CA)--are background assumptions for Dretske. They don't either into the analysis itself; their function is rather to motivate the analysis. Without (IA) and (CA) in the wings, the project to give an account of the causal role of content in folk-psychological explanation degenerates into mere ordinary language philosophy or folk-wisdom journalism.

I think Dretske's motivating assumptions are on the wrong track, or at least seriously misleading. The explanatory role of semantic content in most contemporary cognitive science is not causal and has little to do with the explanatory role of reasons in folk-psychology. In order to substantiate this charge--or at least make it intelligible and plausible--I will need to contrast my view of things with Dretske's. This means that I will have to begin with some critical exposition of Dretske's theory.

Contents as Causes

Exposition

Behavior. Dretske distinguishes the outputs (e.g., movements) of a system from its behavior. Behavior is the production of output by a cause internal to the system. Behavior is thus a process--the causing of M. by C--rather than an event. It follows from this conception that to explain behavior is to explain something of the form C-->M (C's causing M). On Dretske's picture, then, semantic content enters into a causal explanation of behavior when we can explain why C causes M by appeal to the semantic content of C.

Structuring Causes. To understand why C's cause M's in S, we need to know something about the way S is structured. To understand, for example, why the bending of a bimetallic element in a thermostat causes the furnace to go on or off, we need to know that the bimetallic element is a switch that closes (opens) a circuit by bending into (away from) a contact point. By a structuring cause of C-->M, Dretske means a cause of S's having whatever structure underwrites the C to M connection in S. A change in the room temperature can explain why the

bimetallic bent, hence why the furnace went on, but it cannot explain the system's behavior, for that requires explaining why a bending of the bimetallic element in the thermostat causes the furnace to go on.6

Putting these last two points together we arrive at the following result: the semantic content of C enters into a causal explanation of the behavior C-->M of S when C's having a certain semantic content is a structuring cause of C-->M. The semantic content of C has to be what allows us to understand the C to M connection in S.

Indication and Representation. How could the semantic content of C help us to understand the C to M connection in S? Dretske's answer is that in some systems there is a C to M connection because of what C indicates: if C's cause M's in s because the occurrence of C's in S covaries with the occurrence of F's (in the environment or elsewhere in S), then, according to Dretske, (i) in S, C's have the function of indicating F's, hence (ii) C's are representations in s of F's, and so we can say (iii) that we have a C to M connection in S because C has the semantic content it does.

Learning. So the question of the explanatory role of content boils down to this: under what conditions, if any, do we have a C to M connection in S because of C's indicative powers? This happens, according to Dretske, when and only when M is learned as a response to F.

Let's begin with the 'when' part. Suppose S responds to F's with M's. Then s must have some way of detecting the occurrence of F's. That is, there must be some internal state C of S the occurrence of which covaries with the occurrence of F's. C's can then be recruited (Dretske's term) as specific causes of M's. Since learning establishes a C to M connection--i.e., establishes a structure that underwrites a C to M connection in S--and since this happens because C's covary with F's, learning gives us just what the doctor ordered, viz., a case in which it is C's meaning that accounts for (is a structuring cause of) the C to M connection in S, i.e., of S's behavior.

Now for the 'only when' part. Consider a bit of unlearned behavior, i.e., a case in which the C to M connection in S is innate. How could C's content enter into the explanation of the C to M connection in S? There seems only one possibility: Organisms with a C to M connection were selected for because C's indicate (or indicated at one time) F's and the capacity to respond to F's with M's conferred a selective advantage on S's ancestors. In an earlier work,7 Dretske thought that this kind of selectional explanation should be regarded as a case of casual explanation of behavior by content. In Explaining Behavior, however, he rejects this view.

> A selectional explanation of behavior is no more an explanation of an individual organism's behavior--why this (or indeed any) moth takes a nosedive when a bat is closing in--then is a selectional account of the antisocial behavior of prison inmates an explanation of why Lefty forges checks, Harry robs banks, and Moe steals cars. The fact that we imprison people who forge checks, steal cars, and rob banks does not explain why the people in prison do these things.8

The point of the passage is to argue that selectional explanations don't explain why C's cause M's, they only explain why individuals with the C to M connection occur (and, perhaps, predominate) in the current population.9

Summary Statement. Putting all this together, we have the following.

> (1) S's behaviors are processes consisting of the production of an output M by a cause C internal to S.
>
> (2) The semantic content of C's explains why C's cause M's in S when C's

having the content it does is a structuring cause of the C to M connection in S--i.e., when C's having the content it does explains why S has the structure that underwrites the C to M connection.

(3) C's having the content F (C's expressing the property F) is a structuring cause of the C to M connection in S when C's cause M's because it is a function of C's to indicate F's. When C's cause M's in S because it is a function of C's to indicate F's, C's are said to represent F's.

(4) There is a C to M connection in S because C's indicate F's when and only when M is learned as a response to F, the underlying mechanism being the recruitment of C's as causes of M's.

Consequences

Several consequences of Dretske's view are worth emphasizing. They are these:

(1) Only learned behaviors have explanations in terms of semantic content;

(2) not even learned behaviors can be said to be explained by representations;

(3) it isn't current meaning that explains current behavior;

(4) whether current behavior is properly explained by the semantic content of an internal state depends essentially on the system's history.

(1) Only learned behaviors have explanations in terms of semantic content. It is worth pointing out that this runs directly counter to a central empirical claim of most cognitive science of the last twenty years, namely the claim that a great deal of learning is based on innate knowledge: learning itself (some of it) is unlearned behavior explained in terms of unlearned knowledge Dretske's rather surprising denial of this central claim is a direct consequence of his assumption that the explanatory role of content is its role in the causation of behavior (identified earlier as motivating assumption (CA)), together with the idea that behavior is a process consisting of the production of output by an internal cause. There is really only one way in which something x can be said to cause C's to cause M's in S: x restructures (or constrains) the system S in such a way that the occurrence of a C leads to the occurrence of an M. (If something jams the door, then the occurrence of smoke inside is going to make you come out a window.) But if x restructures S so that C's cause M's, then the C to M connection in S is acquired. Unacquired C to M connections cannot be caused to occur because you cannot cause the occurrence of what is already there. Dretske cannot allow for the explanation of unlearned behavior (e.g. learning itself) in terms of innate knowledge.

(2) Not even learned behaviors can be said to be explained by representations. On Dretske's account, C's acquire the function of indicating F precisely by being recruited to cause M's because of the value of outputing M when F.

> Once C is recruited as a cause of M--and recruited as a cause of M because of what it indicates about F--C acquires, thereby, the function of indicating F. C acquires its semantics, a genuine meaning, at the very moment when a component of its natural meaning (the fact that it indicates F) acquires an explanatory relevance.10

It is thus trivial to say that C causes M because of what C represents, for C's status as a representation is constituted by the fact that its (past) indicative power is relevant to explaining the C to M connection. What makes it C's function to indicate F (what makes C a representation

of F) is just that C causes M because of what C indicated during learning.

> A belief is merely an indicator whose natural meaning has been converted into a form of non-natural meaning by being given a job to do in the explanation of behavior.11

It is thus natural meaning that does the explaining, that has been given a job to do in the explanation of behavior. Non-natural meaning can't explain behavior because what makes it non-natural is, by definition, just the fact that the correlative natural meaning explains behavior.

(3) It isn't current meaning that explains current behavior. Once the system is structured in a way that guarantees a C to M connection, C's will continue to cause M's regardless of what C's indicate. (As Dretske is at pains to point out, it can be a function of C's to indicate F's even if C's do not indicate F's.) The structuring cause of the C to M connection is the fact that C's indicated F's during learning. The fact that it is now a function of C's to indicate F's--the fact that C's now represent F's is not a structuring cause of the C to M connection, hence does not enter into the explanation of the behaviors that are constituted by the production of M's by C's.

(4) Whether current behavior is properly explained by the semantic content of an internal state depends essentially on the system's history. Dretske's account shares a disturbing feature with that of Millikan,12 viz., that a just created molecule by molecule duplicate of me cannot behave for reasons. This is because, in the duplicate, the behaviors aren't learned but, as it were, preprogrammed.13 On this conception of things, most of artificial intelligence is based on a conceptual error, for AI assumes that one can simply give an artificial system what natural systems have to learn, and one can do this without copying the physical structure of any natural system.

Two Pictures.

Attractive as Dretske's picture is in many ways, it has, as we've just seen, some unattractive consequences. Three seem especially worth reiterating.

> (1) The account is incompatible with the plausible claim that much learning is dependent on innate knowledge.
>
> (2) The account is incompatible with the plausible claim that cognitive states can be synchronically specified.14
>
> (3) The account is incompatible with the widespread claim that it is current representations and their current semantic contents that (in part) explain current behavior.15

These consequences are disturbing enough, I think, to lead us to ask whether and how they might be avoided.

There are two strategies a revisionist might employ, depending on how the trouble is diagnosed. The conservative diagnosis assumes that the trouble is basically vibration that can be cured by careful tuning and tinkering. The radical diagnosis assumes that the trouble doesn't derive from the details, but from the assumptions that motivate the project--the assumptions (IA) and (CA) identified earlier. In the remainder of this paper, I want to explore a radical alternative to the picture that emerges if one begins with (IA) and (CA).

The Causal Assumption

> (CA) Semantic content cannot have a serious explanatory role to play in

psychology unless it has a causal role to play.

Remember functionalism? Functionalism was the idea that mental states could be individuated via their roles in the causation of behavior or rather, as Dretske has taught us to say, of output. what happens when a functionalist applies this treatment to belief and desire, i.e.c to intentional states? Well, intentional states are mental states that are individuated by their semantic properties, i.e., by their contents. Two things seem to follow. First, semantic properties of mental states must somehow be a matter of causal roles: same causal role, same state (functionalism); different semantic properties, different state (individuation of belief); hence, distinct beliefs must differ in their causal roles. Second, functionalism is plausible to the extent that we think that what matters about mental states--what makes them important--is their role in the causation of output. From this point of view, talk of mental states gets into psychology at all--gets past the behaviorist gate keepers--because mental states are deemed important to the explanation of output. It is still the behaviorist game--explain output--but mentalists win because it turns out that one can't explain output without mental states.

Putting these two points together what we have is this: the game is to construct causal explanations of output by reference to mental states.16 Since what is distinctive about some mental states, namely the intentional states, is their contents, the game in the case of intentional states is to discover the causal role of content: if intentionality is to have a serious explanatory role to play, it must be the case that some states cause output in virtue of their semantic properties. We thus arrive at the causal assumption, viz., that the explanatory role of content is its causal role in the production of output.

This may not be the way Dretske got to the causal assumption, but it is a pretty common and natural way to go for a philosopher influenced by functionalism or a psychologist influenced by the behaviorist definition of psychology as a science in the business of explaining/predicting behavior. Natural as the causal assumption is, however, there are other ways to go, as we'll see shortly.

The Intentionalist Assumption

> (IA) the role of semantic content in folk-psychological explanation just is, or is our best line on, the explanatory role of semantic content in scientific cognitive psychology.

Two facts should make this assumption look dubious to us.

First, the idea that belief and desire involves representation is a daring and controversial empirical hypothesis, an hypothesis I call the representational theory of intentionality. According to this hypothesis, championed by Fodor, to harbor a belief that Brutus had flat feet is a two part affair. It is (i) to harbor a representation that means that Brutus had flat feet, and (ii) for that representation to be (computationally) available to the system as a premise in reasoning, and to be subject to evidential assessment. Representations, in this picture, are conceived as data structures, or something very like them, like them, things that can be 'read' and 'written' in the computational senses of those terms. There is thus a large gap between the role of intentional states (belief, desire) and the role of representation, and a correspondingly large assumption is being made when it is assumed that the explanatory role of intentional state contents is a good guide to the explanatory role of the semantic contents of mental representations.

Second, talk of mental representation, and hence of the relevance of semantic content to

psychology, got a serious scientific start in psycho-linguistics, and in the computational modeling of reasoning. In both cases, the contents of the representations invoked were not plausible candidates for the contents of intentional states (by which I mean ordinary propositional attitudes such as belief and desire), a point misguided critics were at pains to make at the time. Representational psychology (as opposed to intentionalist psychology) didn't start out with belief and desire and branch out into phrase structure and goal hierarchies; it went the other way around.

Given these facts, we should, as I said, be suspicious of the intentionalist assumption. I think suspicion is rare in this context because the rules set by the intentionalist assumption seem (to philosophers) to define the only game in town. But there is another game in town.

The Interpretationist Picture

The computational theory of cognition (CTC) seeks to explain cognition by appeal to two correlative explanatory primitives, representation and computation. It is central to this approach that these are assumed to be well understood in virtue of uncontroversial applications in the explanation of non-cognitive capacities such as calculation and elementary character manipulation. Essential to the appeal of the CTC is the idea that the very same elementary processes that account for the arithmetic capacities of a calculator and the character manipulating capacities of a word-processor can be made to account for sophisticated cognitive capacities as well. For the CTC, it is elementary data structures -- e.g., the stored value of a variable -- rather than beliefs, that are the paradigm cases of representation. Advocates of the CTC reject what I have called the intentionalist assumption and are thus free to remain agnostic about the representational theory of intentionality.17

The CTC takes as its paradigm of the explanatory role of representation not the explanation of behavior by reasons (which may not involve representation at all!), but the computational explanation of calculation and symbol manipulation. To understand the picture from the perspective of a theoretical framework that rejects both the intentionalist assumption and (as we'll see shortly) the causal assumption, we do well to begin by examining the explanatory role representation in elementary calculators and other symbols manipulators.

Adding machines. To add is to compute the plus function. But + is a function on numbers, and numbers are not physical states of calculators. How, then, can calculators traffic in numbers? The answer is simple and familiar: adding machines instantiate + by computing representations of its values from representations of its arguments. A typical adding machine computes a numeral for the number seven from a pair of numerals for the numbers five and two. The numerals in question are physical states of the machine--relative cog wheel positions, say, or relative dc levels. In figure 1, the points along the double line represent physical states of the machine, points on the single line above represent the numbers which are the interpretations of the corresponding physical states. What makes a physical state a numeral is just that it is properly interpreted as a number. A physical state is properly interpreted as a number just in case the diagram commutes, i.e., just in case the conjunction of A_1 and A_2 causes S iff $+(I(A_1,A_2)=I(S)$.

Figure 1: an adding machine satisfying the function
g instantiates the function +.

Typically, of course, the mapping from numeral to number is made obvious by the labels on the keys, and by the use of a display featuring numerals in some standard notation. This makes it easy to miss the distinction between what the states of the machine represent in the machine (if anything) and what the states of the machine mean to us. But there is a difference: whether or not the states of the system are really numerals in the system, and which numbers they represent, depends on whether we have properly designed the system so that its states really do represent what we intend them to (our intentions being recorded as our choice of labels).

That there is a fact of the matter concerning what is represented that is independent of standard conventions or the intentions of designers or users is made clearer by a different sort of example, viz., Galileo's discovery that the elements of geometrical figures represent mechanical magnitudes. Consider a body uniformly accelerated from rest that travels a fixed time t. When time runs out, it will have achieved a velocity v. Now consider a body that travels at a uniform velocity v/2 for the same time t. It turns out that both bodies will cover the same distance. Galileo's proof of this result involves a revolutionary use of geometry. In figure two, the height of the triangle/rectangle represents the time t. The base of the triangle represents the terminal velocity v of the uniformly accelerated object, and hence the base of the rectangle represents the constant velocity v/2 of the unaccelerated object. The area of the rectangle represents the distance traveled by the unaccelerated object (vt), and the area of the triangle represents the distance traveled by the accelerated body.18 Proof of the result reduces to the trivial demonstration that the triangle and the rectangle have the same area.

The crucial point is that, given Galileo's interpretation of the lines and volumes, the laws of Euclidean geometry discipline those representations in a way that mirrors the way the laws of mechanics discipline the represented magnitudes: the geometrical discipline mirrors the natural discipline of the domain. That is, geometrical relationships among the symbols have counterparts in the natural relations among mechanical variables19 in such a way that computational transformations on the symbols track natural transformations of the system.20 This is what makes it correct to say that the symbols - lines and volumes - *represent* times, velocities and distances.

Of course, this is what Galileo *intended* them to represent: that is the interpretation he stipulated. But it's one thing to intend to represent something, another to succeed. Galileo's figures *actually do* represent mechanical variables because the computational discipline *actually*

does track the natural one - the natural discipline we have in mind when we say the system behaves according to natural law. Galileo's interpretation is a *proper interpretation* because, under that interpretation, the natural system and the geometrical system that represents it march in step: the geometrical system *simulates* the natural one. Representation, in this context, is simply a convenient way of talking about an aspect of more-or-less successful simulation. The volumes behave in the geometrical system in a way analogous to the way certain distances behave in the natural one. Hence, the volumes are said to represent those distances; those distances are proper interpretations of those volumes. For instance, the volume of the triangle tracks the distance traveled by the uniformly accelerated body; the volume of the triangle is the *geometrical analogue* of the distance traveled by the accelerated body. This is what makes it correct to say that the volume of the triangle *represents* the distance traveled by the accelerated body, i.e., that the distance traveled by that body is a proper interpretation of that volume. Representation enters into this story in a way exactly analogous to the way it enters into the story about adding machines. In both cases, it is the fact that one function simulates the other under a fixed interpretation that makes it possible to think of the arguments and values of one function as representing the arguments and values of the other. The causal structure of an adding machine - the fact that it executes an appropriate program and hence satisfies the function g - guarantees that the arguments and values of g track the numbers; guarantees, for example, that "3" is the computational analogue (in the machine) of three in the addition function. This is what makes it possible to think of "3" as a symbol *in the system* for three. Analogously, the formal structure of Euclidian geometry guarantees that the volume of the rectangle will track the distance traveled by the unaccelerated body, and this is what makes it possible to think of that volume as representing that distance.21

The important point about this example is that it brings out the fact that the notion of representation we are investigating is no more arbitrary or 'imposed' than the use of mathematics in science generally.22 Galileo's discovery counts as a discovery, not an invention, because his interpretation is a proper interpretation: nature disciplines the relationships between the mechanical magnitudes in question in a way that is mirrored by the way the laws of Geometry discipline the corresponding elements of the figure. Galileo makes an historic contribution by discovering that the formal structure of geometry (properly interpreted) represents the structure of nature. Science has never been the same since. His task, far from being one of arbitrary imposition of interpretations, was to discover whether and how geometrical relations represent mechanical relations.

Every programmer will recognize Galileo's problem. When programming, it isn't enough to baptize one's data structures (or use a natural language); you have to write a program that imposes a discipline on the symbols that mirrors the discipline nature imposes on the things one hopes to symbolize. When you get the algorithm wrong, the representations fail to represent what you intend them to represent.23 Galileo had this problem himself: Euclidian geometry doesn't quite do the job for mechanics, so the representation is imperfect.

The Explanatory role of S-Representation

I propose to call the sort of representation I've been discussing 'S-representation' to distinguish it from other sorts of representation, and to emphasize that its essential feature is that there be a kind of simulation relation between the formal structure of the representational and the

natural structure of the domain represented.24 It should be obvious by now that the explanatory role of s-representation isn't anything like the explanatory role Dretske assigns to mental meaning. The idea behind s-representation is rather that, under proper interpretation what was seen as mere computation (or any other discipline embedding the symbols) is revealed as something else: adding, chess playing, reasoning.

What makes this idea appealing in its application to cognition is just the old idea that cognitive behavior--rational behavior in some sense--is epistemically assessable behavior. The difference between cognizing an environment and simply responding to it is that cognitive behavior is behavior that satisfies epistemic constraints.25 But epistemic constraints are defined over propositional contents. Thus, to be capable of epistemic constraint satisfaction, and hence of cognition, a system must be capable of states that have propositional contents. Cognizers are thus systems that have propositional contents. Cognizers are thus systems whose states have proper propositional interpretations. The CTC's central idea is that it is computational structure (dispositions to compute) that provides the relevant discipline on the states. The hypothesis is that there is an interpretation that will reveal the right cognitive structure. It is an hypothesis analogous to Galileo's hypothesis that there is an interpretation that will reveal geometry as mechanics. The jury is still out on the CTC's hypothesis.

So s-representation is just a name for what you've got when there is a proper interpretation linking two structures/disciplines. Interpretation (and hence representation) contributes to understanding by effecting what Haugeland calls a dimension shift: the hope behind the CTC is that mere computation (of just the right kind) will be revealed by proper interpretation as thinking.

Semantic interpretation is just a special case of the sort of redescriptive conceptual filtering that good scientific taxonomy accomplishes generally. Good scientific taxonomy allows one to describe a situation or domain in a way that filters out everything except the information relevant to the explanatory problem at hand. Equipped with the glasses of newtonian mechanics, what one sees when one looks at an otherwise colorful and complex scene from The Hustler is a plane normal to g, (the surface of the pool table) populated by various vectors (momenta originating at the centers of gravity of the balls). Similarly, equipped with the glasses of proper interpretation, what one sees when one looks at the otherwise complex transactions in an adding machine is calculation. The latter is no more in the eye of the beholder than the former.

Conclusion.

Dretske's target is the role of content in the commonsense explanation of behavior by reasons. It might, therefore, seem unfair to tax him with failure to be compatible with the special empirical assumptions of theories in artificial intelligence and cognitive psychology. But Dretske's project degenerates into mere ordinary language philosophy or folk-wisdom journalism unless he accepts what I have called the intentionalist assumption, the assumption that the ultimate scientific relevance of semantic content to our understanding of the mental is best seen in commonsense explanation of behavior by reasons. I have tried to argue that the intentionalist assumption is by no means inevitable (or even plausible), and that there are other routes into the problem of mental semantics, viz., the use made of it by actual science. Pursuing this routes leads me to reject Dretske's other motivating assumption, the causal assumption, according to which the explanatory role of mental content is to be found in its contribution to the causation of

behavior. Instead, I find that semantics enters into current scientific attempts to understand the mind as bridge over the gap produced by what Haugeland has called a dimension shift.

I have, of course, ignored a major aspect of Dretske's motivation, which is to give a naturalistic account of 'original intentionality',26 i.e., of the ordinary propositional attitudes--belief, desire and so on. But representations and their semantic properties were not introduced into the science of the mind to account for belief and desire, they were introduced to account for the capacity to solve problems, and to parse speech; to account for psychological phenomena such as the Sternberg effect,27 the 'chunking' effect,28 the verbal transformation effect,29 or the scanning effect.30 Indeed, there is no commonsense reason to suppose that belief and desire involve representation at all, and some commonsense reasons to deny.31 Perhaps propositionally interpretable states computationally available as representations of premises for a reasoning algorithm aren't 'real' beliefs, and perhaps propositionally interpretable GOAL states aren't 'real' desires. But they seem to be just what the doctor ordered. I don't know what 'real' belief and desires are, but I have yet to see any compelling reason why a serious science of the mind should care.

References

Cummins (1988) *Meaning and Mental Representation.* Bradford Books/MIT Press.

Dennett (1978) *Brainstorms*. Bradford Books/MIT Press.

Dretske, F. (1988) Explaining Behaivor. Bradford Books/MIT Press.

Dretske (1987) "Misrepresentation," in *Belief*, R. Bogdan, ed. Oxford University Press.

Fodor (1987) *Psychosemantics*. Bradford Books/MIT Press.

Haugeland (1978) "The Nature and Plausibility of Cognitivism," *Behavioral and Brain Sciences,* 2:215-60.

Haugeland (1985) *Artificial Intelligence: the Very Idea*. Bradford Books/MIT Press.

Koslyn, S., T. Ball and B. Reiser (1978) "Visual Images Preserve Metric Spatial Information: Evidence form Studies of Image Scanning," *Journal of Experimental Psychology: Human Perception and Performance*, 4:46-60.

Lewontin, R. (1983) "Darwin's Revolution," *New York Review of Books*, 30:21-27.

Miller, G. (1956) "The Magic Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *Psychological Review*, 63:81-97.

Millikan (1984) *Language, Thought and Other Biological Categories.* Bradford Books/MIT Press.

Pollock, J. (1988) "How to Build A Person," *Philosophical Perspectives*, v. 1.

Sober (1984) *The Nature of Selection.* Bradford Books/MIT Press.

Sternberg, S. (1969) "Mental Processes Revealed by Reaction Time Experiments," *American Scientist*, 57:421-57.

Stich (1983) *From Folk Psychology to Cognitive Science: The Case Against Belief*. Bradford Books/MIT Press.

Warren, R. and R. Gregory (1958) "An Auditory Analogue of the Visual Reversible Figure," *American Journal of Psychology*, 71:612-13.

.1 Fred Dretske, Explaining Behaivor (MIT Press--Bradford Books, Cambridge, Mass., 1988).

. Stephen Stich, From Folk Psychology to Cognitive Science: The Case Against Belief (MIT Press--Bradford Books, Cambridge, Mass., 1983).

. Stich's challenge can be generalized to constitute a challenge to any framework that takes representations to be realized in a medium whose relevant causal properties are presumed independent of the semantic properties of the representations realized.

. Jerry Fodor, Psychosemantics (MIT Press--Bradford Books, Cambridge, Mass., 1987).

Dretske, Explaining Behavior, pp. 79-80.

C-->M is a process. If we think of a process as a causal chain, then one way to causally explain a process is to cite a cause of it's initial event. This is what Dretske calls a triggering cause of C-->M. If B's cause C's in S, and a B occurred in S at t, then the process C-->M will be initiated at t. Triggering causes explain why the C-->M process occurred when it did, but do not explain the C to M connection in S. That requires a handle on the structure of S.

Fred Dretske, 'Misrepresentation,' in Belief, ed. R. Bogdan, (Oxford, Oxford University Press, 1987), pp~~.

Dretske, Explaining Behavior, p. 95.

Dretske (following Sober, following Lewontin) distinguishes selectional from developmental explanations. See Elliot Sober, The Nature of Selection (MIT Press--Bradford Books, Cambridge, Mass., 1984), and R. Lewontin, 'Darwin's revolution', New York Review of Books, 30 (1983), pp. 21-7.

Dretske, Explaining Behavior, p. 84.

Dretske, Explaining Behavior, p. 84.

Ruth Millikan, Language, Thought and Other Biological Categories (MIT Press--Bradford Books, Cambridge, Mass., 1984).

For the careful: On Dretske's account, as on Millikan's, only an accidental duplicate fails to represent. A copy--a duplicate made by copying the original--does have its structure because of the learning history of the original, though it has no learning history of its own.

Taken together, points (ii)-(iv) have the consequences that minds, conceived as systems whose states are, in part, semantically individuated, are very peculiar systems in that their theoretically relevant states do not supervene on even the entire current physical state of the universe. This follows from the fact that physical states are a-historically conceived, while representational states are, according to Dretske, essentially individuated by actual learning histories.

Cognitive state should be distinguished from epistemic or justificatory state: whether I am currently justified in representing the world the way I do doubtless depends on my history. But whether I currently represent the world as, say, containing black swans does not seem to depend on my history. Russell's suggestion that we might, for all we know, have been created five minutes ago, having just the beliefs we now have, doesn't appear to be trivially dismissable on the grounds that beliefs and so on are historically individuated.

For an argument that indicator semantics of the sort that Dretske favors is also incompatible with some fundamental empirical claims of cognitive science, see Robert Cummins, Meaning and Mental Representation (MIT Press--Bradford Books, Cambridge, Mass., 1988), chapter 6.

This looks fishy already because functionalists define mental states in terms of their role in the causation of output, so it seems circular to causally explain output by reference to mental states as causes. What happened to Hume's sound insistence on independent access to cause and effect?

If, as some suppose, intentionality presupposes rationality, those who accept the intentionalist assumption are bound to wind up thinking mental representation requires

rationality too. The CTC is not burdened with this implausible result.

The area of the triangle, of course, is half the base times the height: vt/4. To see that this is the distance traveled by the accelerated object requires some mathematical reasoning that was not formulated explicitly until the invention of the integral calculus.

By mechanical variables, here, I mean real mechanical properties that vary in magnitude. I do NOT mean symbols.

The tracking referred to here is not causal, of course. A computational system can simulate a natural one without there being any significant causal relations between a symbol and the property it tracks in the simulated system. This is important because it allows for the fact that a computational system can simulate hypothetical systems and counter-factual systems, as well as abstract systems and systems that are actual and concrete but not in any significant causal interaction with the simulator.

Notice that we are not talking about a particular distance here - three meters, say - but *whatever* distance an unaccelerated body travels for an arbitrarily specified velocity and time. Plug a velocity and time into the geometry as the base and height of the rectangle respectively, and the volume is the distance traveled.

Dretske thinks that cases of this sort are cases of derived intentionality, i.e., cases in which what meaning there is there is meaning only <u>for us</u>, the effect of convention and or arbitrary imposition.

> Let this dime on the table be Oscar Robertson, let this nickel (heads uppermost) be Kareem Abdul-Jabbar, and let this nickel (tails uppermost) by the opposing center. These pieces of popcorn are the other players, and this glass is the basket. With this bit of stage setting I can now, by moving coins and popcorn around on the table, represent the positions and movement of these players. I can use these objects to describe a basketball play I once witnessed. (Dretske, <u>Explaining Behavior,</u> pp.52-3)

If only Galileo had known it was this easy!

They always manage to represent something, of course: take them as numerals and there is some arithmetic function the program computes (though probably not a familiar or useful one). This bothers those who accept the intentionalist assumption because their paradigm is belief, and beliefs aren't always about numbers as well as people places and things.

Notice that s-representations cannot come one at a time.

John Haugeland, 'The nature and plausibility of cognitivism', Behavioral and Brain Sciences, 2 (1978), pp. 215-60.

John Haugeland, Artificial intelligence: The very idea, (MIT Press--Bradford Books, Cambridge, Mass., 1985).

S. Sternberg, 'Mental processes revealed in reaction time experiments', American Scientist, 57 (1969), pp. 421-57.

George Miller, 'The magic number seven, plus or minus two: some limits on our capacity for processing information', Psychological Review, 63 (1956), pp. 81-97.

R. Warren and R. Gregory, 'An auditory analogue of the visual reversible figure', American Journal of psychology, 71 (1958), pp. 612-13.

S. Koslyn, T. Ball and B. Reiser, 'Visual images preserve metric spatial information: evidence from studies of image scanning', Journal of Experimental Psychology: Human Perception and Performance, 4 (1978), pp. 46-60.

Daniel Dennett, Brainstorms (MIT Press--Bradford Books, Cambridge, Mass., 1978), and, more recently, Dennett, The intentional stance, (MIT Press--Bradford Books, Cambridge, Mass., 1988).