

Teleosemantics and Why Millikan's Theory Eliminates Fodor's Circularity Problem

KENNETH E. STALZER

THOUGHTS AND BELIEFS that we hold are commonly about things; for example, the belief 'that P' is about P. Our beliefs have intentionality, or "aboutness," "of-ness," or "directedness" upon objects. These inner mental states are representations, but at the same time are physical states of the brain. The question then arises: Internal mental representations, the thought 'that P' or the belief 'that P', must get their semantics somehow, but how?

Two philosophers who have attempted to tackle that question from a naturalistic point of view by appealing to a teleological theory are Jerry Fodor and Ruth Millikan. Yet their theories have important differences. The purpose of this paper is to first explicate Fodor's theory and then to argue that there are some circularity problems with it. I will then move on to explicate Millikan's theory, showing how her theory successfully accounts for Fodor's failings.

I. Fodor's Content-Determining Conditions¹

One way to talk of the semanticity of mental representations is to use the term 'truth-conditions'.² Doing so avoids the semantic pitfall of appealing to the intentionality of a thought. Put formally: "S is the truth

Ken Stalzer is a junior in philosophy at the University of Notre Dame. He plans on pursuing a Ph.D. in philosophy, probably specializing in the philosophy of mind.

¹This particular theory of Fodor's was taken from a paper he only reluctantly allowed published, and is not the theory he currently endorses.

²There is a whole theory of truth to which there is a very good introduction in Quine's *The Philosophy of Logic*, but I will attempt to give a cursory

condition of M," where M is a mental representation and S is the actual or possible state of affairs in virtue of which M has its truth value. Here the content of M does not determine its meaning.

Although truth conditions will be necessary to Fodor's theory, they are not sufficient; there needs to be something more. Why is this so? Because there are multiple ways of mapping symbols onto the world. Fodor writes, "There is only one pairing whose extension is the (ordered) pairs of symbols and their truth conditions, there is nothing about the notion of correspondence *as such* which distinguishes this pairing from the others" (*Psychosemantics* 317).

Fodor calls on another condition or relation that will be coextensive with the relation that a truth condition specifies. Like Hume's "M pictures S" is coextensive with 'S is the truth condition of M' and Skinner's "'S is the discriminative stimulus of M' is coextensive with 'S is the truth condition of M,'" Fodor has his own relation R^* which does the work of the first phrase in both cases. The teleological part comes in because Fodor defines this relation based on the

explanation here. A truth theory allows one to go from axioms about references to statements of truth-conditions of a language, i.e. one can postulate a proposition and get truth out of its reference. In Fodor's case, we are talking about the language of thought. Let us take an example. 'Cicero orates' is true if and only if Cicero orates. Now we take each word independently. 'Cicero' [refers to, stands for, or denotes] Cicero. 'Orates' can be either (1) true of x iff x orates, or (2) 'Orates' [refers to, stands for, or denotes] the property of x being an orator, orating etc. Here we are taking axioms about reference moving towards statements of truth-conditions. Let's continue. The Syntax: 'Cicero' is a name [noun phrase, NP]. 'Orates' is a verb [verb phrase, VP]. Rule: x consists of a NP + VP $\rightarrow x$ is a sentence. Now we have a semantic rule for sentences: (1) A sentence NP + VP is true $\leftrightarrow$ the reference of [NP] has reference of [VP] (i.e., has the property of VP), and (2) A sentence NP + VP is True $\leftrightarrow$ the [VP] is true of the reference of the [NP]. So we can replace semantic rule (1) with: 'Cicero orates' is true $\leftrightarrow$ Cicero has the property of being an orator; 'Cicero orates' is T $\leftrightarrow$ Cicero orates; 'P' is T $\leftrightarrow$ P. Similarly for semantic rule (2): 'Cicero orates' is T $\leftrightarrow$ 'orates' is T of Cicero; 'Cicero orates' is T $\leftrightarrow$ Cicero orates; 'P' is T $\leftrightarrow$ P. Therefore the semantic rule yields for each S (sentence) in the L (Language) 'P' is T $\leftrightarrow$ P. This is the accepted theory Fodor is attempting to use.

functional and teleological properties of cognitive systems, as will be seen later.

Now for this relation R^* . Fodor's computational theory of belief states: (a) O believes S if and only if there is mental representation M such that (1) S is the truth condition of M and (2) O bears R^* to M (*Psychosemantics* 319).

Now to avoid circularity in a semantic theory, we cannot appeal to the semanticity of "is the truth condition of," but we can appeal to R^* . R^* is going to be our determining factor in deciding what truth conditions map onto the world. Fodor continues by adding the nice little heuristic of the yes-box. Let's pretend everyone has a box in his or her head that contains beliefs as token mental representations. We'll call it the yes-box. Its contents are determined by this principle: For each mental representation M, O bears R^* to M (and hence believes what M expresses) if and only if a token of M is in the yes box. Hence, relation R^* indicates there is a token of a belief in the yes-box, and moreover that the belief is therefore believed. So what determines what is in the yes-box? Well certainly the yes-box does, but so too does perception, memory, inference—whatever the correct psychology says is involved in the fixation of belief. A belief being in the yes box means something has played this causal role for the belief. For later purposes we'll call this something a *representation producer*.

Now the teleological claims start to come in. Fodor assumes that there are cognitive mechanisms whose function is to put tokens into the yes-box. Fodor writes:

I'm going to specify a certain (nonsemantic) property P such that, or so I'll claim, it's precisely because a mental representation has P that it has the truth conditions it does. Now, P is a property that mental representations have in virtue of facts about the teleology of cognitive systems; mental representations have P because the cognitive systems of organisms have certain normal functions. (*Psychosemantics* 323)

So mental representations enter the yes-box because our cognitive systems were designed to prefer certain Ms to others due to selection pressures as told by a Darwinian story. Now an organism does not want random Ms selected for its yes-box, but rather ones that affect its behavior in a manner

that gives the organism an advantage in survival. Fodor calls such constraints 'entry conditions'. Fodor states, "The entry condition for a mental representation M is that state of affairs such that, when functioning normally, (the cognitive system brings it about that M is in the yes-box) iff (the entry condition is satisfied)" (*Psychosemantics* 324).

It appears that the entry condition is the P Fodor refers to in the quote above. This is apparent because cognitive systems respect these conditions when they are doing what they were selected to do. Cognitive systems, surely, are designed to bring about coherent relationships between inner representations and states of the outer environment. An example of this is the frog's cognitive relationship to a fly passing by: the frog recognizes the fly passing by and shoots out its tongue to catch it.

Because entry conditions are external to the organism, how does the organism decide which external thing to take a representation to be? Fodor answers this question by introducing "conditions of epistemic appropriateness." This will restrict tokening to particular mental symbols. Conglomerating all of this together, Fodor defines his 'entry condition function' thus:

Each mental symbol M gets correlated with the state of affairs S such that: (given the satisfaction of relevant epistemic appropriateness conditions) *it is constitutive of the normal function of the cognitive system that it puts a token of M into the yes box iff S obtains.* (*Psychosemantics* 326)

We see that Fodor is focusing on representation production; that is, when cognitive mechanisms are functioning normally under epistemically appropriate conditions (e.g. enough light to see, not too far away, etc.), then their teleological function is to put Ms into the yes-box. So, assuming we know all there is to know about the teleology of a visual system, and the epistemically appropriate conditions, then the representation under the entry condition function is if 'there is a fly flying around near a frog' then there is a fly flying around for a frog.

Now Fodor comes to the conclusion alluded to earlier, namely that the entry condition for a system of mental representations is coextensive with the truth conditions thereof. That is, S is the truth condition of M if and only if S is the entry condition for M. Fodor writes, "A cognitive system is simply a mechanism whose (appropriately idealized)

input/output function is a truth definition for the representational system that it employs" (*Psychosemantics* 326).

Fodor offers a nice summation of his theory by offering this hypothetical situation. Pretend you are a neuropsychologist and you are attempting to figure out what certain neural spikes represent. Of course, you would try to find the entry conditions of those spikes. Fodor states:

This means (1) you would look for a state of affairs such that bringing it about is necessary and sufficient for the occurrence of the spike given that the organism's visual apparatus is intact and that the relation of the organism to that state of affairs is epistemically appropriate to the visual fixation of belief; and (2) you would ask yourself whether making the assumption that that state of affairs is the truth condition of the spike lets you tell a reasonable story about the behavioral (etc.) consequences of the neuron's firing. (*Psychosemantics* 327)

So, for a frog, the fly dancing around causes the neuron to fire, which causes the tongue to catch the fly, etc.

Fodor's theory has been unpacked. Although it is a good teleological account of how inner mental representations get their semantics, there remain some problems.

II. Fodor's Ideal Conditional Circularity Problem

Fodor's theory is appealing, but it is ultimately circular. Perhaps the main reason is because he focuses on representation producers and their ideal conditions. Let us remember that representation producers are those mechanisms (perception, memory, inference) which aid in belief fixation. Fodor's theory, explained in the previous section, is that to find out the semanticity of representations one must look at what the representations correspond to when their producing mechanisms operate under biologically ideal conditions. Further, there is a concern as to how we are supposed to understand ideal conditions and "epistemic appropriateness." That is, do these conditions specify what 'X' means or is supposed to mean? Fodor includes these conditions in his theory to explain cognitive mistakes, but they end up determining content (as will be explained below).

Unfortunately, the emphasis is on the wrong part of the story, and in Fodor's theory this leads to circularity. Here's how. On his above theory, Xs cause 'X's under ideal conditions, where the cognitive mechanisms are performing according to their design under *epistemically appropriate conditions*. Thus the representation 'X' is *therefore true*—remember the theory of truth conditions. (Further recall that Fodor is focusing on representation producers; this will be key later on). But why are we to suppose the cognitive mechanisms are designed to always yield truths? One quick example: Many women do not remember clearly the pain of giving birth after it is over. Presumably, natural selection designed this mechanism of forgetting to encourage continued propagation of the species.³

But how is Fodor's theory supposed to account for the times when Ys cause 'X's? This is the problem of misrepresentation or disjunction: under ideal conditions cognitive mechanisms are supposed to yield truth, but do not. That is, Xs are the cause of 'X's if and only if Xs are the state of affairs that obtain in the world, but sometimes Ys are really the cause of 'X's.

Let us take an example. Driving late at night, I see a black thing in front of me which appears to have eyes, and I believe that it is a black cat in front of my car. Yet in reality it is a black cloth with shining beads sown on it that appear as eyes. In my yes-box I have the belief that a black cat is in front of my car. We want the symbol in the belief box, 'cat', to be caused by there being *a cat*. We want the symbol to mean what put it in the yes-box. Yet in this case a black cloth kicked 'a cat' into my belief box. Now the black cloth means 'cloth' as well as 'cat'. At this point we cannot appeal to "epistemic appropriateness" or "ideal conditions" because we are using these concepts not to explain what the black cloth means, i.e. 'cat', but rather what it is supposed to mean, i.e., 'black cloth'. Fodor puts in "optimal circumstances," "ideal conditions,"

³Stich writes, "For natural selection does not care about truth; it cares only about reproductive success....What we have shown then is that one inferential system may have a higher level of external fitness than another even though the latter, less fit system makes fewer mistakes and gets the right answer more often" (62). Stich includes other arguments that evolution does not always necessarily select cognitive mechanisms that yield truths under ideal conditions.

and “functioning properly” not to answer in the positive how we get content, but to explain mistakes. But in fact, when they explain mistakes, then *they are answering how we get content*. If we are to determine the semantics of a mental representation, M, by looking at optimal conditions for M’s tokening, we must realize that the optimal conditions for tokening M1 could be different from tokening M2. Only by knowing the content could we figure out the optimal conditions, but the content was what we were trying to figure out in the first place. And so we have circularity, which Fodor himself recognizes:

This raises the possibility that if we’re to say which conditions are optimal for the fixation of belief, we’ll have to know what the content of the belief is—what it’s a belief *about*. Our explication of content would then require a notion of optimality whose explication in turn requires a notion of content. (“Where Do Truth Conditions Come From?” 106)

So, the circularity in Fodor’s above theory has been unveiled. Again, notice that it is Fodor’s emphasis on the optimal conditions for representation *production* that leads him into such problems. I will now explain Millikan’s theory and show how it solves such problems.

III. Millikan’s Teleological-Representation Consumer Theory

Millikan also develops a teleological theory. Although it is teleological, the unpacking of her theory is different than Fodor’s. Fodor’s, remember, was teleological with respect to how the cognitive mechanism produced representations under content-determinate conditions. Instead of looking at these conditions, Millikan argues that “indicating correctly [representing] is a biological *norm for our cognitive perceptual systems...[representing] in the way that has accounted for their survival in the species*” (*White Queen* 7, emphasis added). Her theory is easier to understand once some basic concepts she uses as a foundation are understood. These concepts are “proper function,” “Normal explanation or condition,” “representation producer,” and “representation consumer.” I will treat each separately, and then show how they work together.

A healthy heart pumps blood. Now, immediately intuition tells you that the function of the heart is to pump blood. But surely water

pumps or hand pumps can pump blood as well, and they are not hearts. Millikan distinctly defines "proper function" within biological categories to make her understanding of it explicit. Millikan writes, "Having a proper function is a matter of having been 'designed to' or of being 'supposed to' (impersonal) perform a certain function...and this has to do not with its powers but with its history," (*Language* 17) "functions that have helped account for the survival and proliferation of its ancestors" (*White Queen* 56). So the proper function of a heart is not determined by looking at its presently performed functions. Rather, to discern its proper function one must look at what it has done throughout its evolutionary history and how this has helped the organism survive.

Now to "Normal explanation" and "Normal condition." A Normal explanation, Millikan writes, "explains the performance of a particular function, telling how it was (typically) historically performed on those (perhaps rare) occasions when it was properly performed" ("Biosemantics" 246). Notice that the normal explanation derives not from the present correct functioning of, say a heart, but how its function historically occurred. Normal conditions are the conditions that are stated in giving a full Normal explanation of a proper function. It is imperative to understand that "Normal conditions [you might say 'historically optimal' conditions] for the performance of a certain function...are not at all the same as average conditions under which members of that family have existed" (*Language* 34). Millikan often uses the example of sperm. A very small percentage of the sperm reach the ovum and only one fertilizes the ovum out of the millions whose function is to do so. So, Normal should not be construed as statistically average, but rather in regard to the biological history of what the mechanism is supposed to do.

Now that we have these two concepts in mind, we can move on to the cognitive mechanisms of representation production and consumption.

Millikan breaks up the representational system into three parts: the representation producers, representation consumers, and the representation itself. We'll mainly be concerned with details of the first two. The representation producer can be perception, memory, inference, etc. As I've said earlier, this is what Fodor focused on in order to give him the semantics of mental representations. Millikan focuses instead on the consumer, "what it is to use a thing as a representation" ("Biosemantics")

246). The consumer could be the desire, or what causes the behavior of the organism. Although Millikan focuses on the consumer, all three parts of her theory are of equal importance and must work together (*White Queen* 125–26).

Now let's see how focusing on the representation consumer will allow us to determine the content and semanticity of internal mental representations. Of course, Millikan is going to use the teleology of proper function and normal explanation/condition in her theory. Millikan writes:

Representations are intentional icons the mapping values of the referents of elements of which are supposed to be identified by the cooperating interpreter [consumer]....[T]he Normal explanation of how the icon adapts the interpreter device such that it can perform its proper functions makes reference to the fact that the icon maps onto something else in accordance with a specific mapping function. (*Language* 96, 99)

We can begin to see that by looking at the representation interpreter/consumer in its proper function of mapping onto the world, we see how the representation gets content. "So," Millikan writes, "there must be something about the consumer that *constitutes* its taking the signs to indicate, say, *p*, *q*, and *r* rather than *s*, *t*, and *u*. But if we know what constitutes the consumer's taking a sign to indicate *p*, what *q*, what *r*, etc., then...we can construct a semantics for the consumer's language" ("Biosemantics" 247). This being the case, we can bring in our teleological factors. To find out what constitutes the consumer's mapping of signs onto the world in a certain way, we try to understand proper functions and how these functions are performed normally. Godfrey-Smith, putting it another way, writes:

The content of an inner icon is determined by an interpretation rule that relates the specific form of the inner icon to conditions in the world. This interpretation rule is determined as follows. The consumer mechanisms have been selected to react in a specific way to the form of icons produced, in accordance with what we might call "consumption rules" (not to be confused with interpretation rules). This way of consuming icons has been successful,

because of a certain relation between the state of the world and the ways icons affect action. This set of consumption rules coordinates actions with the world in a way that was selectively successful.
(181)

The desires of an organism which lead to behavior are determined by the representation consumer's mapping the representation on the world and acting in a way that is selectively successful in continuing existence. The content, though, is determined when the cognitive mechanisms are performing their proper functions under normal conditions. It is only then that a correspondence or mapping is reached between the representation and the world.

It would be salient to now turn to some examples. The paradigmatic example in this case is bee dances. When bees return from gathering nectar from flowers they perform a dance "telling" the other bees where to go to find the nectar. The dance "talks" by having the bees dance at different tempos, waggle different ways, have the angle of the dance at different angles to the sun and so forth. The bees who do the dance can be considered representation producers, the dance itself is the icon or representation, and the bees who perceive the dance and then fly away for the nectar could be considered representation consumers. This entire procedure has a complex historical, evolutionary explanation for the interworking of producer, representation, and consumer. The dance is a spatio-temporal map that indicates the location of nectar. Since it is about where nectar is, you might say it has intentionality. The dancing bees have some evolutionary account that explains the mechanisms inside that cause them to dance to indicate where nectar is. Therefore, they have a proper function that is historically and evolutionarily explainable.

The proper function of the watcher bee, then, is to get itself to the nectar. This proper function, getting to the nectar, is realized only if something else holds, viz., if the icon or representation (dance) is mapped to the nectar. The normal explanation of how the watcher bees can perform their proper function, i.e. getting to nectar, only holds when the dance maps nectar accurately. Only in this way are the watcher bees working properly. Therefore, relating the mapping to proper function then enables us to recognize the semanticity of 'nectar at place x'.

A few more examples: A basic survival activity for a species is avoiding danger. Ground squirrels emit a squeal to indicate nearby danger and use different pitches to indicate whether that danger is in the form of a hawk or snake. The shriek that indicates a hawk causes the squirrels to take cover in the nearest brush to avoid the hawk. If it is the shriek that indicates a snake, the ground squirrels know that hiding in the brush would be futile so they must search for the nearest underground hole that the snake could not fit into. When there really is a hawk or snake then the shriek of Squirrel 1, which is the representation producer, excites the sensory apparatus of Squirrel 2. The representation consumer (inferential mechanism, desires, etc.) of Squirrel 2 processes the information, has a desire to survive, and acts accordingly. Squirrel 2 maps representations onto the state of affairs of the world because its cognitive system is acting out its proper function under normal conditions.

It is important to note that on this account it is not necessary to assume that most representations are true. Squirrel 1 may have shrieked because of a kite or model airplane, and Squirrel 2 still would have reacted. Remember that our mechanisms may not fix true beliefs usually or even on average, but just often enough. This is why Normal explanations are derived from the history of the organism. Surely, it is better to have a fallible belief-fixing mechanism than none at all. Millikan writes, "It is actually advantageous to fix too many beliefs, letting some of these be false, rather than fix too few beliefs" ("Biosemantics" 249).

How do these representations, therefore, get their content and semantics? The rule which determines the representations, e.g., the high pitched shriek, is based on the evolutionary history of the organism. The history of the organism defines the relation between the representation and the world, and the representation achieves meaning in causing the organism to act in a selectively successful way. Thus, the ability of the squirrel to recognize the danger signal aids in its survival. Here we are coming to the end of my explanation of Millikan's theory of how internal representations get their semantics. I end with a quote from Millikan:

Representations that the brain manipulates or calculates over are symbols at all...only insofar as they are, first, semantic items, items

that map onto the world when they succeed in performing their full proper functions Normally. (*White Queen* 81-82)

IV. Millikan's Solution to Fodor's Circularity Problem

I have been emphasizing Fodor's reliance on representation producers and his "ideal conditions": it should be obvious that Millikan's focus falls upon representation consumers and incorporates the notion of "ideal conditions" as biological norms for our cognitive mechanisms. Briefly, let me restate Fodor's circularity problem. Fodor hopes to identify optimal conditions for representation production, and if the cognitive mechanisms are working as they were designed then they will get the content of a representation and its meaning correct. The problem with this way of thinking is that optimal conditions can be different for tokening different mental symbols. Therefore, to know what conditions are optimal you need to know what the content of the belief is, but that's what the optimal conditions are attempting to answer in the first place.

Millikan anticipates such problems and focuses instead on the representation consumers and different normative constraints. This should be obvious by the explication of her theory, but why exactly does this solve the circularity problem? We look at the representation consumers and see what effect they produce as their function. We are only concerned, in discovering the semantics, with the Normal conditions for proper performance of the representation consumer.

Let's look at an example. A representation consumer of a frog, when all initial conditions are correct, would bring it about that if a bug was in its visual space close enough to catch then the frog would shoot out its tongue and catch the bug. Now what if a BB was in the frog's visual space, and the consumer enacted that same response. Here is where Fodor runs into trouble, because he is looking at representation-producing conditions to determine the semantics. Millikan, on the other hand, states, "Normal condition for proper performance' of the consumer is, ...in the frog case, that the firing of the detector is coincident with the presence of an edible bug. None of the other correspondences...is relevant to this kind of explanation of the consumer's performance, hence none is relevant to the semantics of the inner representations consumed" (*White Queen* 127). A frog's representation

consumer's Normal explanation of its biological history states that it is not functioning properly under normal conditions when it catches a BB and not a bug. Hence, this misrepresentation is irrelevant because we are not looking at the representation consumer functioning properly under Normal conditions. In Millikan's account her use of "optimal" or "ideal" is taken to mean in a historical, evolutionary account of survival, not simply in the conditions which are best to determine content, and therefore Millikan avoids the circularity problem of Fodor. We see clearly how her focus on the representation consumer and its proper functioning under normal conditions is essential to alleviating the problems with Fodor's theory.

In order to develop a naturalistic account of the semantics of inner mental representations, an appeal to teleology must be made. Specifically, the appeal to teleology should be made in reference to biological conditions and historical evolutionary explanations. Both philosophers have given it a shot, and it appears that Millikan has come closer to the target.

Works Cited

- Fodor, Jerry. *Psychosemantics*. Cambridge, MA: MIT Press, 1987.
- . "Psychosemantics or: Where Do Truth Conditions Come From?" *Mind and Cognition: A Reader*. Ed. William G. Lycan. Oxford: Basil Blackwell, 1990.
- Godfrey-Smith, Peter. *Complexity and the Function of Mind in Nature*. New York: Cambridge UP, 1996.
- Millikan, Ruth. *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press, 1984.
- . *White Queen Psychology and Other Essays for Alice*. Cambridge, MA: MIT Press, 1993.
- . "Biosemantics." *Mental Representation*. Ed. Stephen P. Stich and Ted A. Warfield. Oxford: Blackwell, 1994.
- Stich, Stephen P. *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. Cambridge, MA: MIT Press, 1990.