Millikan.pdf

6

Compare and Contrast Dretske, Fodor, and Millikan on Teleosemantics

By "teleosemantics" is meant a teleofunctional account of what determines the semantic contents of inner representations.[1] One contrast among these three authors is that Millikan and Dretske adopt teleological accounts while Fodor rejects teleosemantics. But I can compare the teleosemantic view Fodor would have held had he not thought better of it, namely, the view he rejects in *Psychosemantics* (1987) and in *A theory of content* (1990). I will emphasize Millikan's view because it seems to be the most difficult—it has at least managed to trip some very competent commentators—and because I have studied it the hardest.

A central problem that teleosemantics is designed to solve concerns mental *mis*representation. It is not possible to accomplish a naturalistic analysis of the representation-represented relation by a single step describing only the relation that holds between the thinker's current states and dispositions and the thinker's environment when she or he harbors a *true* representation. This is because there are two kinds of nontrue representations to be accounted for: those that are not true and those that are not representations. The failure to hold of a single-step true-representation-to-world relation could account for only one of these kinds of failure. To understand false as well as true representation, apparently we must understand what *bare* representation is, and then what being true or false is, over and above bare representation. The problem for a naturalist is to do this without introducing ad hoc abstract objects, say unanalyzed meanings, senses, propositions, or possible states of affairs, as somehow ingredient in nature.

1. Thanks to Steven Wagner for the neologism "teleosemantics."

The teleosemanticist solves this problem by introducing teleofunctions into the analysis. On the dominant current analyses, these are functions that are not built out of current properties and dispositions but rest on historical relations. To have a teleofunction is to have emerged from a certain sort of history, one involving some form of selection. Because of this history, the teleofunctional item counts as being "designed to," or even less formally, "supposed to," have a certain structure, and as being "supposed to" perform a certain function. "If language device tokens and mental intentional states (believing that, intending to, hoping that) are members of proper function or 'biological' [i.e., teleofunctional] categories, then they are language devices or intentional states [representations] not by virtue of their powers but by virtue of what they are supposed to be able to do yet perhaps cannot do. For example, just as hearts and kidneys are sometimes diseased or malformed, so sentences and beliefs are sometimes false, and words and concepts are sometimes ambiguous and sometimes vacuous" (*LTOBC*, 17). The contrasts among the teleosemanticists concern how to employ the notion of a teleofunction in order to yield the notion of a representation of a state that is "supposed to" correspond in a certain way to the environment, even though it may not in fact correspond.

The first disagreement among our three authors concerns which are the mechanisms whose functions are relevant to mental semantics. According to Fodor's original plan, it is the representation-producing mechanisms that should be examined. To discover what the semantic content is of the representations these mechanisms produce, we should ask to what these representations correspond when the producing mechanisms operate under biologically ideal conditions. For example, granted that the conditions for perception are ideal, we should expect "That's a horse" to be tokened in Mentalese when and only when that's a horse. The problem with this suggestion, as Fodor notes in *Psychosemantics*, is that ideal conditions for making one sort of observation are not the same as ideal conditions for making another. For example, ideal conditions for seeing very small objects are different from ideal conditions for seeing very large objects. In general, what constitute ideal conditions for making a particular judgment depend upon the content of the judgment. The content of the judgment cannot, then, be derived from the ideal conditions without circularity. If the teleofunction of the representation producing devices is the relevant factor to explore, it seems that we must address this function more directly.

Dretske's teleosemantic proposal is to do just that. There are biological systems, he claims, that have as their functions to represent or "indicate" states of the environment. On a selectionist account of function, this means that they have been selected during evolutionary history "*because* they played a vital information-gathering role . . . essential . . . to the satisfaction of a biological need," or it means that they have been selected during a learning process for a similar reason.[2] A problem with this suggestion, one that Dretske highlights in "Misrepresentation" (1986) but soft pedals into a footnote in *Explaining behavior* (1988, 63), is that for typical cases it seems to yield *numerous* representeds corresponding to each inner representation. Dretske's key example concerns the magnetotactic systems of certain Northern Hemisphere bacteria. These systems contain tiny magnets that pull toward magnetic north, hence toward geomagnetic north, hence down, hence away from the surface of the water, hence toward regions of lesser oxygen, oxygen being toxic to these organisms. Which of these various things is it the function of the magnetosome to indicate? Dretske claims that the answer is indeterminate. For example, "this primitive sensory mechanism is, after all, functioning perfectly well when, under [a] bar magnet's influence, it leads its possessor into a toxic environment" (1986, 29), because it still correctly indicates magnetic north. Fodor's example of this same problem concerns the frog's (neural) fly detector. Surely, he claims, this mechanism is designed to detect small ambient black things and small shadows crossing the retina as much as to detect flies.[3]

As a corrective to the emphasis that others in the teleosemantic business have placed on the function of the representation producers, Millikan (chapter 4 herein) has recently been emphasizing the devices that use or "consume" representations. The official statement of Millikan's position, *LTOBC*, however, emphasizes producer and consumer equally. It also distinguishes the functions of these two from that of a third and quite different thing, the representation itself. The roles that these

2. Caution: Dretske (1988) claims to be agnostic on the question concerning the philosophical analysis of biological function talk.
3. Fodor (1990) claims it is also designed to detect flies-or-BBs, but that is a different matter. On this, see Millikan 1991a.

three items play are distinct but equally important for an analysis of mental semantics.

The indeterminancy problem that Dretske and Fodor encountered is solved by examining the role of the representation consumer with care. Representation consumers are devices that have been designed (in the first instances, at least) by a selection process to cooperate with a certain representation producer. The producer likewise has been designed to match the consumer. What the consumer's function is, what it is supposed to effect in responding as it does to the representations it consumes, could be anything at all. It may have numerous alternative functions. It may also be but one of many consumer systems that use representations made by the same producer. The consumer operates, of course, *after* the producer does, and a *full* explanation of how the consumer has historically managed to perform its function or functions—in Millikan's terminology, a full "Normal explanation for proper performance" of its function—would include that the *producer* first performed *its* function properly, and it would include an explanation of *how* the producer's function has historically been accomplished. A "most proximal Normal explanation for proper performance" of the consumer's function, on the other hand, is one that does not go into any events that occur in the production-consumption chain prior to the consumer's activities. It begins at the point at which the consumer enters the event chain and explains how, given certain initial surrounding conditions, given the consumer's normal structure or constitution, and given certain initial relations between consumer and environment, the consumer has historically produced the effects that are its functions. It explains in the briefest possible way that is still *complete* how the consumer does this for *just* those (possibly infrequent) cases when the consumer has succeeded. That is, the point is not to explain the frequency or infrequency of the consumer's successes, but to explain success in just those cases in which success occurs.

Now *that* sort of explanation will necessarily make reference to an initial relation among consumer, representation, and environment that is of interest to the teleosemanticist. For surely one relation that is strikingly implicated is the *coincidence* between the representation that the consumer confronts and a certain condition in the environment, namely, the condition that, consequent to the analysis, will be designated as

the one the representation represents. Prior to analysis, this condition is picked out in accordance with a certain *rule* of correspondence having the following property: unless we assume that some actual condition in the world corresponds in accordance with this rule to the representation confronted by the consumer, we cannot account, *with any single explanation that covers historical instances of consumer successes generally*, for *why* the consumer produces the effect that is its function. In the case of the magnetosome, this crucial initial relation (this particular "Normal condition for proper performance" of the consumer) is that the magnet points toward lesser oxygen; in the frog case, that the firing of the detector is coincident with the presence of an edible bug. None of the other correspondences mentioned, as above, by Dretske or Fodor is relevant to *this* kind of explanation of the consumer's performance, hence none is relevant to the semantics of the inner representations consumed.

But now the question arises, What makes these hypothesized semantically relevant correspondences different from yet other correspondences that appear to figure with equal importance in "most proximate normal explanations" for performances of consumer functions? For example, surely the occurrence of the magnetosome's signal needs to be coincident with the presence of an immediate surround of water if its consumer is to perform properly, and the firing of the frog's bug detector needs to coincide with presence of a surround of air and presence of a stable platform under the snapping frog. This is where the function of the representation *producer* needs to be brought in. The representation producer has been designed by selection to produce representations for the consumer *that* correspond to conditions in the world *by* the rule of correspondence that figures in the most proximate normal explanation of the consumer's successes. To be very explicit, the producer's job is to produce not just a representation—graphically, a "shape"—but to produce a correspondence, a certain relation *between* "shape" and world. Obviously, if this is the producer's function, there must be a way that it sometimes *effects* this function. But the magnetotactic system does not help to *effect* that the tug of the magnet coincides with the presence of a surround of water, nor does the frog's bug-detection system help to *effect* that its firings coincide with the presence of a surround of air and presence of a frog-supporting platform. So it cannot be a *function* of

these systems to produce these correspondences, hence these correspondences are not relevant to mental semantics.[4]

What, then, are the functions of representations themselves? It is not uncommon among teleosemanticists to suggest that a representation is something whose function is to represent, or to "indicate." To claim this is not necessarily to run in circles. A toggle reamer, after all, is something whose function is to ream toggles. But then we had better know what it is to ream a toggle and, similarly, what it is to represent or indicate something. (Let me drop the "or" and just say "indicate.") There are two sensible things that "indicate" might mean in this context.

Indicating might be, at least in part, a standing in a relation of correspondence to or coincidence with something thereby indicated. In that case, indicating would be not something that a representation *effects* but some way that it *is*, namely, standing in a certain relation. But functions are things *effected* by items having functions. To speak otherwise is to confound what is surely best kept separate. Besides functions, there are, after all, normal explanations both of the genesis of various biological items and of their proper operations, and there are normal conditions associated with both of these kinds of normal explanations. By reference to these various categories, everything that needs to be said can surely be said, and said clearly, without having to blunt the term "function." So if indicating is standing in a relation of correspondence, then it is not one of the *functions* of a representation to indicate anything. Rather, it is a function of a representation *producer* to *produce* a representation that indicates.

But there is another thing that indicating might be: it might be something done by the representation to the representation's consumer, namely, the representing of certain conditions *to* that consumer. Think of this act of representing as the effecting of a change of a certain sort in the consumer. In that case, the representation may be seen as having representing as a genuine function. Specifically, to represent circumstance *c* to consumer (interpreter) *i*, the representation effects a change in *i* that adapts *i*'s further activities to *c*, that is, modifies *i*'s activities

4. This is the answer to the puzzle about determinacy, given Millikan's account, which Cummins attempts to solve by introducing what he calls "basic factors" (1989, chap. 7).

so that *i*'s teleofunctions get performed in, or via mediation of, or despite, *c*.

After the smoke has cleared, it appears that Dretske and Millikan agree on at least two general points: that inner representations are produced by systems having as teleofunctions the production of true representations, and that misrepresentation occurs when these systems miss. Yet there is a strong contrast between their views on how to unpack "true representation" in the context of this analysis, that is, on the details of the job description that is written for representation producers.

Dretske begins with a story about natural signs, or "indicators," and then adds teleology to it. His story goes through several versions (1981, 1986, 1988), but should there be some inconsistencies among these, they do not affect the point at issue with Millikan. For Dretske, an "indicator" is not just an item that actually corresponds to or actually coincides with the affair it indicates. It is one the occurrence of which, as a *type*, makes entirely certain, or at the bare minimum highly likely, the existence of an affair of the *type* indicated. For example, the firings of the frog's bug detector cannot be indicators of bugs unless the probability that a bug is present is high given that the detector has fired. The job of a representation producer is to make indicators in this sense of "indicator," according to Dretske. Indicators made by such a producer are full-fledged, true "representations" (Dretske 1988). When the producer's product fails to coincide with what it should indicate, this product is a false "representation." (False representations do not "indicate" in Dretske's terminology.)

A problem with Dretske's view is that it is hard to see how it could be the function of any biological device literally to *effect* the production of one of his "indicators." To do so, the device would have to *effect* that certain statistics should hold. The frog's bug-detecting system, for example, would have to literally *effect* that the statistics on bugs versus BBs in its environment should be such that, when it fires, the chances favor a bug. But *that*, it surely cannot do. All it can *effect* is that coincident with a bug, there happens a firing. It can at most effect first one such coincidence and then later another.[5]

5. Insofar as it manages to effect such coincidences, it will, of course, *alter* the relevant statistics somewhat or, in some cases, greatly. But altering the statistics and effecting them whole are quite different things.

Millikan describes the function of the representation producer accordingly. Its function is to produce representations, "shapes," that *correspond*, that is, correspond in the manner required by the representations' consumers in order that these consumers should function properly. If circumstances are not normal for the producer, that is, if circumstances are not as they have historically been when the producer has succeeded in its task, then, of course, the producer will almost undoubtedly fail. For example, the frog's bug-detecting apparatus fails whenever a BB shadow crosses the frog's retina. It fails to produce a true representation. But this apparatus cannot be held responsible for the likelihood versus unlikelihood of BBs. It cannot be held responsible for the statistics on its *rate* of success or failure.

But, you may object, is it not part of a representation producer's job to be reliable? Surely the bug detector wouldn't have been selected had it been too unreliable. The frog wouldn't then have had *this* kind of bug detector, but another kind. True, but *that* observation can be made of *any* item that has a function. No teleofunctional item is such that it would have been selected if it had been "too unreliable" about the performance of its functions. But this fact cannot turn being reliable into a part of the function of every teleofunctional item. Indeed, to reiterate, being reliable can't be the function of *any* teleofunctional item for the easy reason that no item *effects* its own reliability. Reliability always depends on the dependability of *external* factors, on the prevalence or rarity of normal conditions for proper performance.

There is a second point of importance here too. Consider: how unreliable is "too unreliable"? Sperm tails, as the Millikan litany goes, are overwhelmingly unreliable at performing their function of propelling the sperm to an ovum, but not, apparently, "*too* unreliable." Similarly, the food detectors built into goldfish are not, apparently, "too unreliable," despite the fact that a goldfish in a dirty bowl may spit back out all or nearly all of the particles these detectors have instructed it to ingest. Nor, apparently, are the danger signals that various species of animals employ "too unreliable," despite the fact that many, in all probability, are much more often wrong than right. What counts as "too unreliable" is a function of the costs incurred when representations turn out false versus the gains that are made when they turn out true

and also of the costs when the animal *fails* to signal what should be signaled. These equations have different values for each kind of representation-producing mechanism.[6] Reliability, then, can be no part of the *definition* of representation any more than it is part of the definition of other kinds of teleofunction.

A third contrast among the teleosemanticists is the special emphasis that Millikan alone places upon the *articulateness* of all complete representations. Complete representations represent complete states of affairs. Complete states of affairs are, as such, articulated, though this articulation may sometimes be very simple, as in the case of the state of affairs constituted by a bug's being here now (rather than here then or there then or there now). A representation that represented something simpler than a state of affairs, one that represented, say, only an object or a property or a *type* of state of affairs (compare a propositional function), would make no *claim*, hence would fail to be true or false, to represent anything either correctly or incorrectly. It would be, or be similar to, a name, *saying* nothing at all. As Frege saw, only when placed in a completing context, such that along *with* this context it represents an articulate state of affairs, does a name truly represent anything.

Now it certainly is possible that an unarticulated representation should be *defined* for use on some special occasion to serve as the name of some state of affairs. It could be defined, that is, by someone who employs a prior articulated representation for this purpose. Thus Paul Revere's single lantern once named the state of affairs that was the British approaching on April 18, 1775, by land. But any such representation will have a use only *once* with any one interpreter. When it has been used once, it has been used *up*. A representational system consisting of just one inarticulate representation can represent just one state of affairs; hence it would have no reason to be perpetuated.

On the other hand, suppose that Paul Revere's signal had not inarticulately signaled the coming of the British by land for that day only but had articulately signaled their coming by land "tonight," the date of the signal standing for the date of the coming. In that case, the signal might be usable again, for the British might come again. But, to think more carefully, it would not in fact *be* the same signal that was used again

6. For a discussion of these equations, see Godfrey-Smith 1991, chap. 10.

but another signal, one with a different date, hence one that represented a different state of affairs. An articulate representation belongs to a *system* of representation, which *system* can be perpetuated. Conversely, barring an explicit convention such as Paul Revere's, set up especially for a single anticipated occasion, only an articulate representation can represent what has not been represented before. For example, only an articulate representation, such as one whose time and place represent that same time and place, can represent to the frog that a bug is here now.

Representation articulation, Millikan claims, is at root the same phenomenon as what is more fashionably termed "compositionality." The principle governing it is the principle of "projection" in the mathematical sense. Even the most primitive of representations ("intentional signals" and "intentional icons") are abstract "pictures," in the sense of *Tractatus logico-philosophicus*, or "maps," in the mathematician's sense, of what they represent. The most sophisticated of linguistic representations are articulate, are compositional, in accordance with the same basic principle as are the most primitive, though the rules of projection involved are, of course, far more complex.

The last contrast that I will mention among the teleosemanticists concerns the role that learning plays in defining the functions that determine mental semantics. Fodor does not seriously consider the possibility that representational capacities resulting from learning might be described in teleofunctional terms. He seems to assume that the only functions that could be relevant to mental semantics are functions derived directly from Darwinian natural selection during the evolution of a species. Representations involving concepts that are not innate, he takes it, could not possibly be treated by a teleosemantic analysis. "You can now see why Darwinian/teleological apparatus does no good," for "when the cognitive mechanisms are behaving as the forces of selection intended them to," "there is no Darwinian guarantee that a properly functioning intentional system ipso facto has the [learned] concept HORSE (to say nothing of the concept PROTON)" (Fodor 1987, 116–17).

Dretske (1988) handles this problem by assuming the classic position that learning is parallel to natural selection. He claims that a learning process may result in the selection or "recruitment" of (what was previously only) an internal *natural* indicator of some condition *c* in an

(individual) animal's environment to serve as stimulus for a behavior that brings a reward when performed under *c*. When that happens, a *function* is bestowed on that indicator, the function of indicating. Dretske does not greatly elaborate this theme. I do not think, however, that it would be opposed to the spirit of his enterprise to extend the notion of behavior in this context to cover inner behaviors, such as thinking processes, and to extend the notion of a reward to cover inner rewards, such as the confirmation of one's beliefs or the avoidance of contradiction.

If Dretske were to allow that sort of extension, he would be moving closer to agreement with Millikan, who likened associative learning to natural selection. A special clause is built into her definition of "proper function" (teleofunction) in order to accommodate associative learning as one originator of proper functions (*LTOBC*, 24). Later the processes that result in the formation of theoretical concepts are described as involving something like trial and error learning for those (possibly rare) cases when concepts are not acquired through a public language. The trials and the rewards here are for the most part inner, with the law of contradiction playing a moderately traditional role. But Millikan adds a string to this bow that increases its range considerably. This is the doctrine of derived proper functions (*LTOBC*, chap. 2) in accordance with which certain kinds of teleofunctions that are built into an animal during evolutionary history interact with the environment of the individual animal to produce new teleofunctions, new biological purposes for these individuals, without the mediation of addional selection processes (*LTOBC* and chapters 3, 11 herein). That learning effects the creation of new teleofunctions is due very largely to this factor, even when learning works by trial and error, by generation and test. And, of course, there are important forms of learning that don't work this way.

The doctrine of adapted and derived proper functions has also another use in the explanation of how new mental teleofunctions are born. It explains how the teleofunctions of elements in a public language become translated into teleofunctions attaching to items in individual language learners' heads (*LTOBC*, chap. 9). I should like to say much more about the notions of adapted and derived proper functions, and about the multiple uses of these notions. But the examination hour is over. Perhaps I shall write my dissertation on this.