

Does Representational Content Arise from Biological Function?¹

Richard J. Hall

Michigan State University

Let us assume that some organisms, humans at least and the other higher animals, have internal states and behavioral states that represent things external to themselves. One of the questions that everyone would like answered about these states is: In virtue of what does such a representational state get the specific content that it has? An answer to this question that's popular just now is: In virtue of its biological function. I believe there is a deep reason why such an answer can't work. I shall present that reason in this paper.

First, we need to be clear about the notion of biological function that's being appealed to here. In particular, we need to distinguish two fairly different ways of thinking about function. We could think of something's function as what it contributes to the present functioning or output of some complex system of which it is a part. Roughly, this is Cummins' view in (Cummins 1984). Or we could think of function etiologically and, in the case of biological traits, evolutionarily: what that trait did that led animals with it to be favored by natural selection. Millikan (1984, 1989), Neander (forthcoming), and Papineau (1987) all advocate some version or other of this etiological notion of biological function. Thus, Neander (forthcoming) states that

"...proper functions are selected effects, and, more particularly, ...biological proper functions are effects for which traits were selected by natural selection. ... hearts have the proper function of pumping blood, ...because pumping blood is what hearts did that caused them to be favoured by natural selection."

And Papineau (1987, p.65) says

"I favour a natural selection account of teleology. That is, I think that when we talk of some characteristic C being present *in order to* produce E, we should understand ourselves to be claiming that C is now present because of some past selection process that favoured C because it produced E."

It is this evolutionary notion of function which is deployed in the biological function theory of representation that I will be criticizing. Notice that a trait need not perform its biological function all the time or even very often for it nevertheless to have

that biological function. "A proper [biological] function of the swimming mechanism in a human sperm", says Millikan (1986,p.58), "is to get the sperm to an ovum. But only a minute proportion of these mechanisms actually perform this function — because of the paucity of human ova." Also, the biological function that a trait now has in a species need not be what the trait was first selected for but rather what it was most recently selected for or what sustains it.

(Both Karen Neander (unpublished) and Paul Griffiths (1989) have suggested that the etiological explanation of the prevalence of some trait amongst the entities of some kind doesn't have to involve *biological* evolution; just so long as that trait's producing some effect E confers some tendency to be selected, we will be able to think of E as the trait's function and as the explanation for why that trait exists in the kind in question. For example, even some trait of an artifactual kind might be explained this way if there is a history of human modification of the artifact, with this trait coming to be standard because of humans selecting instances with it in preference to instances without it. In the remainder of this paper, I shall be concentrating exclusively on biological function, and when I use the word 'selection', I mean natural selection.)

So how is it supposed that the content of a representational state arises out of its biological function? In this way. If the (would be) representational state R produces some behavior b, and b works to the evolutionary advantage of the animal when and only when external condition C obtains, then R represents (means, has the content) C.

"The point of emphasizing [behavioral] outputs is that it allows us to fix the truth condition of a representational state as that circumstance in which the behavioural output of the state produces successful results." (Papineau 1987,p.71)

Millikan gives the following example to illustrate this notion:

"...consider beavers, who splash the water smartly with their tails to signal danger. This instinctive behavior has the function of causing other beavers to take cover. The splash means danger, because only when it corresponds to danger does the instinctive response to the splash on the part of the interpreter beavers...serve a purpose. If there is no danger present, the interpreter beavers interrupt their activities uselessly. Hence, that the splash corresponds to danger is a normal condition for proper functioning of the interpreter beavers' instinctive reaction to the splash." (Millikan 1989,p.288)

In this example, the tail slapping is the representational state R, the diving of the other beavers is the behavior b, and the presence of danger is the external condition which is represented.

Putting this biological function view of representation explicitly, we get something like:

(A) R represents F iff the biological function of R is to produce behavior b when F's are present.

In the following discussion I shall concentrate on representations that are fairly simple and direct, fairly "signalish" one might say, cases like Millikan's beaver tail slap above. The reason is two fold: first, these are the cases the biological function advocates often start off with themselves, and second, in these cases the candidates for biological function and for associated external condition are much clearer. Much clearer than what? Much clearer than in typical cases of human representation, e.g.

human belief or desire. It is very hard to see how to give an account of the content of my belief that $2+2=4$ in terms of its biological function and the external conditions necessary for the performance of that biological function. Since I shall be raising doubts about the biological function account in even the simple signalish cases, I assume that a biological function account of *all* representational content including the human belief that $2+2=4$ would be even more difficult to bring off. Concentrating on cases of beavers and other non-human animals raises the question of whether such animals have any representational states in the first place. I think they do, but if one thought they didn't (and yet that humans do), then one should be deeply suspicious of the biological function account of representation. For that account makes essential appeal, in its etiological account of biological function, to evolution, and it is near the core of evolution to see continuity between humans and other animals and to see precursors of human abilities in lower animals.

The biological function view wants to say that the meaning or content of the beaver's tail slap is "Danger" (or more plausibly, "Dangerous predator nearby"), because the biological function of the tail slap is to produce a certain behavior (diving) when dangerous predators are nearby. That is, tail slapping got selected because it produced diving when predators were nearby. At first, this seems plausible. But a trouble lurks just under the surface: The biological function of a state is seriously underdetermined, underdetermined in ways that result in an unacceptable underdetermination of the content of representational states like the beavers' tail slap. Let me explain.

The biological function theorists maintain that biological function is determined by the *actual* evolutionary history of the species, by what actually got the state selected. At first glance this may seem to yield the right result. In the actual evolutionary history of beavers, avoiding predators via tail slap-caused diving is what got tail slapping in the presence of predators selected. The trouble is that in the actual history of the beaver species, not all beaver predators played a role in the selection of the tail slap-dive adaptation. Suppose, for simplicity, that bears and wolverines are the natural predators of beavers. Not all bears and wolverines will have interacted with beavers during the latter's evolutionary history. Only some finite set of them will have. Any finite set of things can be specified in lots of ways, e.g. via a property they uniquely share or by naming the individual members. It just so happens (let us suppose) that the bears which played a role in the evolution of beavers were the bears born when the sun was in Aquarius. Similarly, the wolverines which played a role were those whose birth weight was a prime number of grams. Aquarial bears and primal wolverines, let us call them. Or (to name the members of the set and to keep its imagined size to an unrealistic minimum), Barney the bear and Bo the wolverine. So why isn't the biological function of the beavers' tail slap: causing diving in the presence of Aquarial bears and primal wolverines? Why isn't the biological function of the tail slap: causing diving in the presence of Barney and Bo? And derivatively, why doesn't the tail slap have for its representational content: "Lo, a bear whose sign is Aquarius or wolverine whose birth weight in grams was a prime number"? Or: "Lo, Barney or Bo"?

I strongly suspect the following principle of biological selection is true:

If 'F' and 'G' are extensionally equivalent predicates (in the relevant environments), then behaving in manner b in the presence of F's will get selected if and only if behaving in manner b in the presence of G's gets selected.

This is the principle that evolution is extensional. To paraphrase Jerry Fodor (unpublished), "Darwin cares how many predators you avoid but not what description you avoid them under." If this is correct, then it is hard to see how the biological func-

tion account can account for extensionally equivalent but intensionally different representations. Maybe beavers don't make such fine distinctions, but we humans certainly do. And of course beavers really do too; no right-minded beaver has a representation with the content "Aquarial bear or primal wolverine", or worse, with the content "Barney or Bo".

One might think that appeal to Sober's distinction (Sober 1984, pp.97ff) between selection *for* and selection *of* can save evolution from extensionality.² For selection *for* is not extensional. Sober's "selection toy" gives a neat, non-biological illustration of this (Sober 1984, p.99). It consists of a vertical cylinder with several horizontal dividing partitions between the top and the bottom. Each horizontal dividing partition has holes of a given size in it, the top partition having the biggest holes and each succeeding lower partition having smaller holes. Balls of several different sizes are put in at the top. All but the biggest balls get through the big holes in the top partition. At the next partition, the next biggest balls are stopped. Only the smallest balls make it all the way to the bottom. They are the ones that are "selected". Suppose that the balls of different sizes are colored differently, and suppose in particular that the smallest balls are green. In the "environment" of this selection toy, greenness and smallness are coextensional, so green balls and small balls are "co-selected". But we, knowing how the mechanism works, can say that the device selects *for* smallness, not greenness; since smallness, not greenness, is the causally efficacious property. Thus, selection *for* is not extensional. (Semantical naturalists, wanting to reduce representation to some nonintensional concept, may have qualms about appealing to this notion of "selection *for*"; let us waive those qualms here.)

The nonextensionality of selection *for* can now be carried over to the notions of biological function and representational content by characterizing biological function in terms of selection *for* rather than in terms of plain selection. Applying this to our beaver case, the biological function advocate will point out that we know enough about evolution to know that it is predatoriness, not Aquarial bearness and primal wolverinness, nor identity with Barney or Bo, which is the crucial property, the one avoidance of instances of which by beavers gets selected *for*. So that's what the biological function of the tail slap is. Consequently the tail slap represents predators, not Aquarial bears and primal wolverines or Barney and Bo.

But this won't do. For even if we grant that selection *for* is not extensional and that avoidance of predators, not avoidance of Aquarial bears and primal Wolverines, nor avoidance of Barney and Bo, is what's selected *for* (so that "production of diving in the presence of predators", not "production of diving in the presence of Aquarial bears...", is the biological function of the tail slap), that doesn't show that the beavers' *representation* has to have the content "predator" rather than "Aquarial bear or primal wolverine" or "Barney or Bo". Any representational state which produces predator avoidance behavior in the presence of predators will be selected *for*, even if that representational state has "Aquarial bear or primal wolverine" or "Barney or Bo" for its content. To see this clearly, let me go back to Sober's example and liven it up a little. Imagine that the balls in Sober's example are animate objects and that they can change their size, shape and color (so that size, shape, color become "behavioral" properties, so to speak). In particular, the balls can change their diameter while leaving shape and color unchanged. Naturally, when they diminish their diameters and make themselves smaller, they make themselves more dense. Imagine they have a self-density detector but not a self-size detector. They learn that when they reach a certain density, they pass through the holes (and are selected). We know that size is what's being selected *for* and that the selection mechanism doesn't care about density at all. Our law of evolutionary selection will state that although dense things are se-

lected, it's small things are selected for. But what the balls learn is "Be dense if you want to make it to the bottom" — and what they represent is density, not size.

Consider a more realistic but still Sober-like example. Imagine a species of animal which lives in an environment in which there are both poisonous berries and edible berries. Members of this species will be selected for avoiding the poisonous berries. Imagine that the poisonous berries are red and the edible ones green and that the animals that avoid the poisonous berries do so by virtue of perceiving that they are red and avoiding the red berries. "Darwin cares that you avoid poisonous berries but not what description you avoid them under", Fodor's still small voice (Fodor's still small voice?) can be heard saying once more.

Just in case the reader is not convinced yet that there are real situations like this with locally coextensional evolutionarily relevant properties, let me give one more example. Consider a kind of worm that has evolved in an environment where all and only its predators cast moving shadows. In worm world, "predator" and "casts moving shadows" are extensionally equivalent. Isn't it natural to suppose the worm detects and represents moving shadows even though there's no biological law that the avoidance of moving shadow casters are selected for? Representing and avoiding casters of moving shadows is how the worm achieves the avoidance of predators. The kinds we use in describing evolution don't have to be the kinds the animal uses, so long as they come to the same thing, namely keeping the animal from getting eaten.

Think of evolution as a big Sober-type sorting device which is continually selecting for avoidance of poisonous berries, avoidance of predators, etc. But this evolutionary sorting device doesn't care *how* a species manages to achieve these things. Doing it with a color sorting device or a shadow sorting device is just fine, as long as color or shadow goes along with the important property. Putting all this in an only slightly over-intellectualized way, we can say that if F and G are extensionally equivalent (in the relevant environments), then nature can't distinguish between the following:

- (i) Detection of F's plus following the rule "If F's are present, then do b".
- (ii) Detection of G's plus following the rule "If G's are present, then do b".

("Detection" is taken as intentional here, as involving representations.)

The biological functionist may still argue that the biological function theory of representational content, as given in (A) above, should be accepted, on the grounds that the specification of the biological function and the representational content of an animal's state is done by *us*, with *our* explanatory and predictive interests in mind, and therefore with *our* best choice of predicates for these purposes. It is we who are taking the intentional stance towards the animal and hence we who specify its intentions. The biological function of the beavers' tail slap-dive mechanism is the avoidance of predators, not the avoidance of Aquarial bears, etc., because that's *our* best description of it in the context of evolutionary theory. So the content of the tail slap as given by this biological function is "Predator nearby", not "Aquarial bear ... nearby". That is what is given by the biological function under our best specification of it.

There are two troubles with this move. The first is that even if one is sympathetic to this somewhat anthropocentric, somewhat instrumentalistic attitude, it's not the case that the only considerations relevant to our choice for the content of the representational state are those concerning evolutionary and selection. It may be, for example, that the *physiology* of the relevant perceptual mechanisms dictate conclusively that it

is some other property than predatoriness or poisonousness that the animal is perceiving and responding to, color, say, or size or motion or shadow. These latter then become the preferred candidates for what the animal is representing. The second and deeper trouble is that this move will lead to a bifurcation in the notion of representational content. Our original notion of representational content was connected, at least in our human case and that of higher animals, with what we are consciously aware of. What we are led to on the proposed interpretation of the biological function account is a notion of content that may have no connection (except local coextensiveness) with conscious content. The biological functionist would describe the content of the internal state of the worm which causes it to crawl down its hole when a moving shadow falls across it (= when a predator is nearby) as: "predator nearby". But the worm, if it had conscious states and could talk, might say: "What I was consciously aware of, what I perceived, was a moving shadow. And I believe that its best to crawl down my hole when I perceive a moving shadow." Perhaps worms don't have conscious mental contents, in which case the divergence I am pointing to won't arise in their case. But at some point as we climb up the rungs of the animal kingdom, or rather as we meander through the horizontal maze of animal forms, conscious mental content occurs — at least with us humans if nowhere else. And my point is that there is no guarantee that this conscious mental content will have anything (except local coextensiveness) to do with the representational content now being stipulated by the biological function account.

So either the biological function theory leaves representations too underdetermined to distinguish between extensionally equivalent but intensionally very different contents, or, by tightening up the specification of representational content through using the intensional notion of selection for, it gives wrong specifications, at least if we want this representational content to agree with the content conscious subjects are themselves aware of. It is hard to see how to overcome this difficulty while still remaining in the biological function camp.

Notes

¹I would like to thank Karen Neander for helpful discussions on this topic.

²Or equivalently, appeal to the notions of law and explanation. There is presumably a law, roughly speaking, that behavior which results in the avoidance of predators gets selected, while there is no law that behavior which results in the avoidance of Barney and Bo, or in the avoidance of Aquarius bears and prime wolverines gets selected. We can explain the presence of the tail slap-dive behavior by saying that it was selected *because* it resulted in the avoidance of predators, but if we say beavers do it because it resulted in the avoidance of Barney and Bo, we don't seem to have given an explanation at all.

I believe the appeal to law and explanation is equivalent, in this context, to the appeal to selection for. I believe the following equivalences hold:

- (1) Property F (as opposed to the coextensional property G) is selected for iff having property F explains the selection whereas having property G does not.
- (2) Property F (as opposed to the coextensional property G) is selected for iff there is a law that individuals with F will be selected but no such law for G.

References

- Cummins, R. (1984), *Psychological Explanation*. Cambridge: MIT Press.
- Fodor, J. (unpublished), "A Theory of Content", unpublished manuscript.
- Griffiths, P. (1989), "A Defence of Etiological Theories of Biological Function", paper read at the 1989 annual Australasian Association of Philosophy meetings, Canberra, Australia.
- Millikan, R.G. (1984), *Language, Thought, and Other Biological Categories*. Cambridge: MIT Press.
- (1989), "Biosemantics", *Journal of Philosophy* LXXXVI: 281-297.
- (1986), "Thoughts Without Laws; Cognitive Science With Content", *Philosophical Review* XCV: 47-80.
- McGinn, C. (1989), *Mental Content*. Oxford: Blackwell.
- Neander, K. (forthcoming), "Functions as Selected Effects", *Philosophy of Science*.
- (unpublished), "The Teleological Notion of 'Function'", unpublished manuscript.
- Papineau, D. (1987), *Reality and Representation*. Oxford: Blackwell.
- Sober, E. (1984), *The Nature of Selection*. Cambridge: MIT Press.