THE JOURNAL OF PHILOSOPHY

VOLUME LXXXIV, NO. 4, APRIL 1987

FUNCTIONS

IN describing the function of some biological character, we describe some presently existing item by reference to some future event or state of affairs. So the function of teeth at time t is to pulp food at time t', where $t' > t$. This seems to present an exact parallel to the case of the function of humanmade artifacts; for instance, the function of the nutcracker at time t is to break open nuts at time t', where $t' > t$.

In the case of biological functions and in other cases of functions where human intentions are not obviously causally active, this kind of description of function has been difficult to assimilate into our scientific view of the world. There are several reasons, but we shall here concentrate on one which arises directly from the fact that, in describing a present structure in terms of its function, we mention a future outcome of some sort. The future outcome may be, in many cases, nonexistent. A structure may never be called upon to perform that function. The function of a bee's sting, for instance, is relatively clear; yet most bees never use their stings. Likewise teeth may never pulp food, just as nutcrackers may never crack nuts.

Thus, when we describe the function of something in the present, we make reference to a future event or effect which, in some cases, will never occur. Hence, prima facie, we cannot really be describing any genuine, current property of the character.

I. THE PROBLEM

Even when a character does perform its supposed function, the future events that result from it cannot play any significant "scientific" role in *explaining* the nature and existence of the character. The character has come into existence, and has the properties that it does have, as a result of prior causes. It would still have existed, with just the current properties it does have, even if it had not been followed by the events that constitute the exercise of its alleged

0022-362X/87/8404/0181$01.60

function. Hence its existence and properties do not depend on the exercising of its function. So it is hard to see what explanatory role its functions could have. Crudely put: backwards causation can be ruled out—structures always have prior causes—hence reference to future events is explanatorily redundant. Hence functions are explanatorily redundant.

Of course, there is nothing inappropriate about describing a character *and mentioning* its future effects. But describing a character as having a function is not just mentioning that it has certain effects. Not every effect counts as part of its function. And some functions are present even when there are no relevant effects to be mentioned —as with some bees and their stings. Future events are not unmentionable; but they are explanatorily redundant in characterizing the existence and current properties of a character. Hence, what role can functions have in a purely scientific description of the world: how can they be "placed" within the framework of current science?

There are three main theories that attempt to construe functions in a way that allows them to fit smoothly into the scientific, causal order. We believe that each is nearly right or partly right. Yet they are all unsatisfactory in one crucial respect: they do not restore to functions any significant explanatory power. In particular, they deny to functions any causal efficacy. So, for instance, they will not permit us to explain the evolution of a character by saying that it evolved *because* it serves a specific function.

We will offer an account of biological function and of functions generally which, although it shares much with the most promising extant theories, is crucially different from them in that it bestows greater explanatory power upon functions.[1] But first we will briefly consider the three theories.

II. ELIMINATIVISM

There are three responses that arise naturally in the face of the tension between functions and the scientific standpoint. The first is eliminativist. It is assumed that functions, if there were any, would have to be important, *currently existing,* causally active, and explanatory properties of a character or structure. It is also assumed that functions, if there were any, would essentially involve reference to

[1] Our theory will be a cousin of theories which are sometimes called "goal theories" and which have been advocated, for instance, by Christopher Boorse, "Wright on Functions," *Philosophical Review,* LXXXV, 1 (January 1976): 70–86, and "Health as a Theoretical Concept," *Philosophy of Science,* XLIV, 4 (December 1977): 542–573; and by William Wimsatt, "Teleology and the Logical Structure of Function Statements," *Studies in the History and Philosophy of Science,* III, 1 (May 1972): 1–80. Some key differences between their theories and ours will be noted later.

future, possibly nonexistent, events. Yet something involving essential reference to future, possibly nonexistent events could not possibly characterize currently existing, explanatory properties. It is thus concluded that there really are no functions in nature.

To add functions to the scientific biological picture, on this view, is parallel to adding final causes to physics. Final causes have no place in the scientific account of the physical universe, and, if the psychological pressures are resisted, we find we can do without them and final causes just fade away. To the eliminativist, the same will be true of functions; as the biological sciences develop, any need for function talk will vanish, and the psychological naturalness of such talk will fade away with time and practice.

A variant on this eliminativist view adds an account of why the attributions of function seem to serve a useful purpose in everyday and scientific discourse. The eliminativist does not believe in functions as genuine, currently existing properties of a character. But the eliminativist does believe in future effects of a character. And nothing stops us from mentioning whichever future effects we take an interest in. Consequently, an eliminativist can interpret talk of "functions" as being merely the specification of effects one happens to be interested in. Which effects are deemed to relate to "functions" of a character, will depend not on the nature of the character itself, but on our interests. The function of kidneys is different for the anatomist from what it is for the chef. Insofar as function talk makes sense, it does *not* describe the current nature of a character (there are no functions in nature); rather, it relates a current character to a future outcome, in an interest-dependent, extrinsic manner.[2]

The best answer to an eliminativist theory is to come up with an adequate analysis of functions still within the scientific view. This is what we attempt to do later in this paper.

But a motive for seeking such a noneliminativist account can be cited; and this motive will also provide a less than conclusive, but nevertheless weighty, argument against eliminativism. In the biological sciences, functions are attributed to characters or structures, and these attributions are intended to play an explanatory role which cannot be squared with the eliminativist's account of function talk. For instance it is assumed that biological structures *would* have had the functions they do have even if we had not been here to take an interest in them at all. And some of the effects of structures that we

[2] A theory of this sort can be found in Robert Cummins, "Functional Analysis," this JOURNAL, LXXII, 20 (Nov. 20, 1975): 741–765. We have also been influenced by a paper by Elizabeth W. Prior, "What Is Wrong with Etiological Accounts of Biological Function?," *Pacific Philosophical Quarterly*, LXVI, 3/4 (July/October 1985): 310–328.

take an interest in have nothing to do with their function. And some functions are of no interest to us at all. Furthermore, biology standardly treats function as a central, explanatory concept. None of this rests easily with an eliminativist theory. A powerful motive for resisting eliminativism, then, is that an adequate analysis of functions, if we can find one, will enable us to take much biological science at face value: we will be relieved of the necessity of undertaking a radical reformation of the biological sciences. The eliminativists' vision, of functions "fading away," is as yet just a pipe dream; and their explanation of the apparent usefulness of function talk fails to explain away more than a fragment of the uses functions serve in biological science.

It is not the biological sciences alone which could be cited here. Psychology could be canvassed too. And even physics has facets that raise problems for an eliminativist view. Suppose someone were to suggest that the function of water is to refract light and the function of mists to create rainbows. Presumably this is plainly false. Yet it describes something in terms of future effects in which we take an interest . . . which is exactly what the eliminativist takes to be the business of function talk. So eliminativists have no good explanation of why the physicist takes such attributions of function to be plainly false. They cannot explain the manifest difference between "The function of mists is to make rainbows," and "The function of teeth is to pulp food."[3]

III. REPRESENTATIONAL THEORIES

There is a response to the tension between the scientific view and functions which rejects any role for future events in the characterizing of functions. The future effects of a character do not themselves play an explanatory role in characterizing that character. Yet sometimes there exists, prior to the character, a *plan,* a representation of that character *and* of its future effects. Such a representation of future effects may exist, whether or not those effects ever come to pass. And this representation exists prior to the character and so contributes to the causal processes that bring that character into being—by the usual, forward-looking, causal processes that rest comfortably within our over-all scientific image of the world. On this view, we can account for a function not by direct reference to any

[3] This problem is treated at some length by Ernest Nagel, *The Structure of Science* (New York: Harcourt; London: Routledge, 1961), in his section on teleology, pp. 401–428. Nagel manages to blunt the force of such objections, but only by augmenting his initial theory, which differs from eliminativism only superficially, thereby generating a theory that comes close to the theories of Boorse and Wimsatt mentioned above.

future event, but rather by reference to a past representation of a future event. Theories of this sort have frequently been called "goal theories"; but they are best construed as a subcategory within the class of goal theories. They are goal theories in which the identification of a goal depends on the content of prior representations.[4]

This kind of account of functions fits best with attribution of functions to artifacts. The idea of breaking open nuts seems to have played a causal role in the production of the nutcracker. It does not fit so neatly into the biological sciences. Of course, it used to provide a persuasive argument (the teleological argument) for the existence of a Creator: there *are* functions in nature; functions require prior representations; yet the creatures themselves (even when creatures are involved) have no such foresight or were not around at the right time; hence the prior representations must have been lodged in some awfully impressive being . . . etc. Nowadays, however, in the clear and noncontroversial cases of functions in nature, it is taken that they can be accounted for from the standpoint of a theory of evolution by way of *natural* selection—and in such a theory there can be no room for any analysis of biological functions which rests on prior representations. Even if God foresaw the functions of biological structures, that is a matter outside biology; functions, however, are a biological and not a theological matter.

It is worth noting that, even though the representational theory seems to rest comfortably with attributions of functions to artifacts, nevertheless, some artifacts prove more problematic than might first appear. Many artifacts evolve by a process very like natural selection. Variations often occur by chance and result in improved performance. The artisan may not understand fully the reasons why one tool performs better than others. Yet, because it performs well, it may be copied, as exactly as possible. The reproduction of such tools may occur for generations. The features of the tool which make it successful and which lead it to be selected for reproduction are features that have specific functions. But they were not created with those functions in mind. They may have been produced with an over-all function in mind (say, hitting nails); but the toolmaker may not have in mind any functions for the components and features of the tool, which contribute to the over-all function. For instance, the toolmakers may copy a shape that has the function of giving balance to the tool—but they need not foresee, or plan, or represent any

[4] Andrew Woodfield, *Teleology* (New York: Cambridge, 1976), argues for a view that takes the primary cases of functions to rest on a prior plan, and all other cases of (unplanned) "functions" to be mere metaphorical extensions of the primary cases.

such function. They know only that tools like this work well at banging nails, or sawing wood, or whatever the over-all function might be. Consequently, even with artifacts, structures can serve specific functions even though there exists no prior representation of that function.[5]

There is a further reason for uneasiness about the representational theory. The theory analyzes the *apparent* forward directedness of functions by an indirect, two-step, route. The forward directedness of functions is analyzed as comprising a *backward* step to a representation, which in turn has a *forward* directedness toward a possibly nonexistent future state.

Thus the seeming forward directedness of functions is reduced to another sort of forward directedness: that of representations—plans, beliefs, intentions, and so on. And this is worrying. The worry is not just that these are "mentalistic," and just as problematic as functions—and just as hard to assimilate into the scientific picture of the world. Rather, an even greater worry for many will be that of vicious circularity. Many find it plausible that the notion of representation will turn out to be analyzable in terms that at least include functional terms. And functional terms presuppose functions. Hence the future directedness of representations may turn out to presuppose the future directedness of functions. This threatens to do more than just restrict the scope of representational theories; it undermines such theories even in their home territory, as applied to artifacts.

IV. ETIOLOGICAL THEORIES

The third response to the tension between the scientific view of the world and the concept of function again involves rejecting any role for future events in the characterization of functions. Yet etiological theories also eschew reference to prior representations of future effects, as well as reference to the future effects themselves.[6]

Representational theories and etiological theories have an important feature in common. Both shun any genuine, direct reference to

[5] There are several intriguing points made about artifacts, their reproduction, selection, survival, and so forth, by Ruth Millikan, *Language, Thought and Other Biological Categories: New Foundations for Realism* (Cambridge, Mass.: MIT Press, 1984). Millikan advances a sophisticated version of the etiological theory, which we discuss below.

[6] The etiological theory is widely held, but a very good exposition and defense is given by Larry Wright, "Functions," *Philosophical Review,* LXXXII, 2 (April 1973): 139–168; and *Teleological Explanations* (Los Angeles: California UP, 1976). We have also been greatly influenced by the defense of etiological theories advanced by Karen Neander, *Abnormal Psychobiology,* Ph.D. thesis, La Trobe University, 1983.

future effects, and refer instead only to past causes. Both construe the attribution of function as supplying information about the genesis of the character, that is, about how the character came into existence.

The difference between the two theories recalls the distinction Charles Darwin drew between artificial selection and natural selection. When animal breeders select, they represent to themselves the characters they wish to develop. Natural selection has closely analogous results, but it operates in the absence of (or at least without any need for) any conscious or unconscious representations of future effects.

The etiological theory of functions explains biological functions by reference to the process of natural selection. Roughly: a character has a certain function when it has evolved, by natural selection, *because* it has had the effects that constitute the exercise of that function.

Clearly, there is room here for an overarching, disjunctive theory, which unites the representational with the etiological. Such an overlapping theory would say: a character has a certain function when it has been selected because that character has had the relevant effects. In the case of artifacts, the selection involves conscious representations (mostly); in the case of (Darwinian) sexual selection, representations may enter the picture, but reproduction, heredity, and evolution also play a part; and in the case of natural selection, representations drop out altogether.

But, on the etiological theory, a character has a biological function only if that character has been selected for by the process of natural selection because it has had the effects that constitute the exercise of that function. This is the only kind of selection compatible with the dictates of modern biological science.

We take this etiological theory of biological function as the main alternative to the account of biological function we shall proffer. The big plus for the etiological theory is that it makes biological functions genuinely explanatory, and explanatory in a way most comfortable with the modern biological sciences. But we shall argue that this explanatory power is still not quite right: it offers explanations that are too backward-looking. The theory we offer will be more forward-looking in its explanatory nature.

But, before we turn to this matter, we should note another worry with the etiological theory, a worry that extends even to the overarching disjunctive theory of which it is a part. This worry is that there is too great a dependence of the intrinsic nature of functions on contingent matters—matters which, had they been (or even if

they are) otherwise, would rob the theory of any viability as an account of functions.

The etiological theory of biological functions has such functions characterized in terms of evolution by natural selection. Most theorists take evolutionary theory to be true, but contingently so. What if the theory of evolution by natural selection were to be (or had been) false? Clearly then, on the etiological theory of biological functions, as we have specified it, there would be no biological functions. Whether or not there are biological functions at all, on the over-arching, disjunctive theory, will depend on what replaces the theory of evolution by natural selection. Suppose it is creationism. Then the representational theory would apply; for we have the representations in the mind of the creator that would have the appropriate causal role in the development of biological structures, so the representational theory would become a general theory of functions.

We noted earlier that the representational theory had problems with the functions of some artifacts: artifacts that seemed to evolve over time by a process similar to natural selection. If this analogous process is also not available, along with natural selection proper, then our representational theory will not bestow functions upon such artifacts. But perhaps this is small change.

For creationism, there would of course be an enormous epistemological problem of discovering what the functions of biological structures were; for this would depend on discovering what the creator had in mind. So we would be stuck with great difficulty in discovering whether the function of the heart is to produce the sound of a heart beat, in line with the creator's idea of a beating rhythm in nature, with the circulation of the blood as a nonfunctional effect; or whether the reverse is true. It would be much like an anthropologist discovering the nature of ancient artifacts without any presuppositions about the intentions of the earlier cultures.

But suppose creationism is not the alternative. Consider the possible world identical to this one in all matters of laws and particular matters of fact, except that it came into existence by chance (or without cause) five minutes ago. Now, on even the over-arching theory, there are no functions; for there are no biological functions on the etiological theory, and no causally active representations as required by the representational theory, and hence no functions at all.

We have the intuition that the concept of biological function, and views about what functions biological characters have, are not thus contingent upon the acceptance of the theory of evolution by natural selection and on discovering what led to the evolutionary develop-

ment of particular characters. In parallel, we are also inclined to think that the representational account of the function of artifacts gives too much importance to the representations or ideas of the original planner, even in cases when there is one. As we indicated earlier, we believe a satisfactory account of functions in general, and of biological functions in particular, must be more forward-looking. We now turn to this.

V. FITNESS, FUNCTION, AND LOOKING FORWARD

It emerges from our discussions that the tension between functions and modern, causal science has generated, fundamentally, two stances on the nature of functions.

The first is the eliminativist stance. This has the merit of giving full weight to the forward-looking character of functions, by specifying them in terms of future and perhaps nonexistent effects; and also to the explanatory importance of functions. It is mistaken only in its despair of reconciling these two strands.

The second stance is backward-looking. This embraces theories which look back to prior representations and those which look back to a prior history of natural selection and those which look back to a history of either one sort or the other.

We will argue for a forward-looking theory. Functions can be characterized by reference to possibly nonexistent future events. Furthermore, they *should* be characterized that way, because only then will they play the explanatory role they need to play, for instance, in biology. The way to construe functions in a forward-looking manner, we suggest, is (roughly) to construe them in the manner of dispositions. The shift we recommend, in our conception of functions, has a precedent: the analysis of the evolutionary concept of fitness.

One wrongheaded, but at times common, objection to the Darwinian theory of evolution is that its central principle—roughly, "the survival of the fittest"—is an empty tautology which cannot possibly bear the explanatory weight Darwin demands of it.[7] This objection

[7] For those who have suggested this view, with greater or less refinement and sophistication, see J. J. C. Smart, *Philosophy and Scientific Realism* (New York: Random House, 1963), p. 59; H. G. Cannon, *The Evolution of Living Things* (Manchester: University Press, 1958); C. H. Waddington, *The Strategy of the Genes* (London: Allen & Unwin, 1957), pp. 64/5; A. O. Barker, "An Approach to the Theory of Natural Selection," *Philosophy*, XLIV, 170 (October 1969): 271–290; Robert Brandon and John Beatty, "The Propensity Interpretation of 'Fitness': No Interpretation Is No Substitute," *Philosophy of Science*, LI, 2 (June 1984): 342–347. For those who have replied to this view, see in particular Edward Manier, " 'Fitness' and Some Explanatory Patterns in Biology," *Synthese*, XX, 2 (August 1969): 206–218; and Michael Ruse, "Natural Selection in the *Origin of Species*," *Studies in the History and Philosophy of Science*, I (February 1971): 311–351.

assumes that fitness can be judged only retrospectively: that it is only after we have seen which creatures survived that we can judge which were the fittest; moreover, it assumes that the fact that certain creatures have survived, whereas others did not, is what constitutes their being the fittest.

The etiological theory of biological functions rests on the same sort of misconception as that which underlies the vacuity objection to Darwin. On this theory, we can judge only retrospectively that a character has a certain function, when its having had the relevant effect has contributed to survival. Indeed, on the etiological theory, that an effect is part of the function of a character is *constituted* by the fact that having this effect has contributed to the survival of the character and of the organisms that bear it.

Consequently, the notion of function is emptied of much explanatory potential. It is no longer possible to explain why a character has persisted by saying that the character has persisted because it serves a given function. To attempt to use function in that explanatory role, would be *really* to fall into the sort of circularity often alleged (falsely) against the explanatory use of fitness in Darwinism.

This comparison with fitness serves another purpose. It has displayed why functions would lack explanatory power on the etiological theory, but it also shows how to analyze functions so as not to lose this explanatory power. Fitness is not defined retrospectively, in terms of actual survival. It is, roughly, a dispositional property of an individual (or species) in an environment, which bestows on that individual (or species) a certain survival potential or reproductive advantage. This is a subjunctive property: it specifies what will happen or what is likely to happen in the right circumstances, just as fragility is specified in terms of breaking or being likely to break in the right circumstances. And such a subjunctive property supervenes on the morphological characters of the individual (or species).[8] Hence there is no circularity involved in casting fitness in an explanatory role in the Darwinian theory of evolution. In the right circumstances fitness explains actual survival or reproductive advantage, just as in the right circumstances fragility explains actual breaking. In each case the explanation works by indicating that the individual has certain causally active properties that in such circumstances will bring about the phenomena to be explained.

What holds here of fitness holds, too, of biological functions. The

[8] For more on dispositions and their supervenience on categorical bases, see Pargetter and Prior, "The Dispositional and the Categorical," *Pacific Philosophical Quarterly*, LXIII, 4 (October 1982): 366–370; and Prior, Pargetter, and Frank Jackson, "Three Theses about Dispositions," *American Philosophical Quarterly*, XIX, 3 (July 1982): 251–257.

etiological theory is mistaken in defining functions purely retrospectively, in terms of actual survival. Hence there need be no circularity in appealing to the functions a character serves, in explaining the survival of the character. Fitness is forward-looking. Functions should be forward-looking in the same way and, hence, are explanatory in the same way.

VI. THE PROPENSITY THEORY

Here is one way to derive a "forward-looking" theory of functions.

Let us begin with the etiological theory. Consider a case in which some character has a specific effect and has been developed and sustained by natural selection because it had that effect. In such a case the etiological theory deems that it is (now) a function of the character to produce that effect.

Look more closely, then, at the past process that has "conferred" a function, according to the etiological theory. The character in question must have had the relevant effect, on a sufficient number of occasions—and in most cases, this will have been not on randomly chosen occasions, but on *appropriate* occasions, in a sense needing further clarification. (For instance, sweating will have had the effect of cooling the animal—and it will have had this effect on occasions when the animal was hot, not when it was cold.)

The history that confers a function, according to the etiological theory, will thus display a certain pattern. The effect that will eventually be deemed a function must have been occurring in appropriate contexts; that is to say, it must have been occurring in contexts in which it contributes to survival, at least in a statistically significant proportion of cases.

Further, this contribution to survival will not, in realistic cases, have been due to sheer accident. One can imagine individual incidents in which a character contributes to survival by sheer chance. Laws of probability dictate that a long run of such sheer accidents is conceivable; but it will be very unlikely—except for characters with relatively short histories. The only cases in which such long runs are likely to occur are cases in which the character confers a standing propensity upon the creature, a propensity that increases its chances of survival.

If we imagine (or find in the vast biological record) a case in which a character is sustained by a chance sequence of accidents, rather than by a standing propensity, then it would not be appropriate to describe that character as having a function. This can happen when a character is linked with another character that does bestow a propensity and where variations in the character just have not occurred to allow selection against the inoperative character. It can also happen by sheer chance, a long-run sequence of sheer flukes. Such a

sequence is very improbable; but biology offers a stunningly large sample. It is very probable that many improbable events will have occurred in a sample that large.

Consequently, what confers the status of a function is not the sheer fact of survival-due-to-a-character, but rather, survival due to the propensities the character bestows upon the creature.

The etiological theory describes a character *now* as serving a function, when it *did* confer propensities that improved the chances of survival. We suggest that it is appropriate, in such a case, to say that the character *has been serving that function all along*. Even before it had contributed (in an appropriate way) to survival, it had conferred a survival-enhancing propensity on the creature. And to confer such a propensity, we suggest, is what constitutes a function. Something has a (biological) function just when it confers a survival-enhancing propensity on a creature that possesses it.

Four features of this propensity theory of biological functions should be made explicit.

First, like the corresponding account of fitness, this account of functions must be relativized to an environment. A creature may have a high degree of fitness in a specific climate—but a low degree of fitness in another climate. Likewise, a character may confer propensities which are survival-enhancing in the creature's usual habitat, but which would be lethal elsewhere. When we speak of the function of a character, therefore, we mean that the character generates propensities that are survival-enhancing in the creature's natural habitat. There may be room for disagreement about what counts as a creature's "natural habitat"; but this sort of variable parameter is a common feature of many useful scientific concepts.

Ambiguities will arise especially when there is a sudden change in the environment. At first, we will refer the creature's "natural habitat" back to the previous environment. But eventually we will transfer the term to the current environment. The threshold at which we make such a transference will be vague. The notion of natural habitat will also be ambivalent as applied to domestic animals.

In its most obvious use, the term 'habitat' applies to the physical surroundings of a whole organism. But we can also extend its usage, and apply the term 'habitat' to the surroundings of an organ within an organism. Or to the surroundings of a cell within an organ. In each case, the natural habitat of the item in question will be a functioning, healthy, interconnected system of organs or parts of the type usual for the species in question. When some of the organs malfunction, then other organs, which go on performing their natural functions, may no longer be contributing to survival. We still say they are

performing their natural function, even though this does not enhance the chance of survival. Why? Because it would enhance survival if the other organs were performing as they do in healthy individuals.

Consequently, functions can be ascribed to components of an organism, in a descending hierarchy of complexity. We can select a subsystem of the organism, and we can ascribe a function to it when it enhances the chances of survival in the creature's natural habitat. Within this subsystem, there may be a subsubsystem. And this may be said to serve a function if it contributes to the functioning of the system that contains it—provided all other systems are functioning "normally" (that is, provided it is lodged in its own "natural habitat"). And so on.

Similar hierarchies may also occur in the opposite direction: a microscopic organism has a function in a pond, which has a function in a forest, which has a function in the biosphere, and so on.

Secondly, on the propensity theory, functions are truly dispositional in nature. They are specified subjunctively: they *would* give a survival-enhancing propensity to a creature in an appropriate manner, in the creature's natural habitat.[9] This is true even if the creature does not survive or is never in its natural habitat. Likewise, fragility gives a propensity to break to an object in an appropriate manner, in the right circumstances—and of course some fragile objects never break. And fitness gives a propensity to an individual or species to survive, in an appropriate manner, in a specified environment and in a struggle for existence, even if there is no struggle for existence or if the individual or species fails to survive.

Of course, when functions do lead to survival—just as when dispositions are manifested—the cause will be the morphological structural form of the creature and the relationship between this form and the environment. Functions supervene on this in the same way that dispositions supervene on their categorical bases. But the functions will be explanatory of survival, just as dispositions are explanatory of their manifestations; for they will explain survival by pointing to the existence of a character or structure in virtue of which the creature has a propensity to survive.

Thirdly, in the long run, it will be necessary to spell out the notion

[9] It is this central role we give to propensities which distinguishes our theory from others, like those of Boorse and Wimsatt (mentioned above), which fall back on the notion of statistically normal activities within a class of organisms. On their theories, a character has a function for a creature when it *does* help others "of its kind" to survive, in a sufficiently high proportion of cases. On our view, frequencies and statistically normal outcomes will be important evidence for the requisite propensities. But there are many well-known and important ways in which frequencies may fail to match propensities.

of a "survival-enhancing propensity" in formal terms, employing the rigors of the probability calculus. Clearly, there will be a spectrum of theories of this general form. These theories will vary in the way they explicate the notion of "enhancement": whether they construe this as involving increasing the probability of survival above a certain threshold, or simply increasing it significantly above what it would have been, and so on. We are not attempting to find and defend *the* correct propensity theory, but only arguing that a propensity theory offers the most promising theory of functions.

Fourthly, there is the question as to whether the scope of the propensity theory is limited to biological functions or whether it can be extended, in some sense, to artifacts.

Obviously, like the etiological theory, the propensity theory could be part of an overarching, disjunctive theory which analyses biological functions in terms of bestowing propensities for survival and the function of artifacts in terms of prior representations. We noted earlier some problems for a backward-looking theory, even for artifacts. Yet surely representations should have some causal role in the case of consciously produced artifacts.

We are attracted here to a general, overarching theory, but one that concentrates on the *propensity* for selection. So a character or structure has a certain function when it has a propensity for selection in virtue of that character or structure's having the relevant effects. In the case of biological functions, we have a propensity for survival in the natural habitat, and so in such a habitat natural selection is likely to be operative. In the case of artifacts, we have a selection process clearly involving representations. But the representations are those at the time of selection, at the time of bestowing the function: now, so to speak. They need not be blueprints that antedate the first appearance of the prototype. Thus we feel there is a sense in which all functions have a commonness of kind, whether they be of biological characters or of artifacts.

VII. COMPARISONS

We hope to have led our readers to appreciate the attractiveness of the propensity theory of biological functions (and of functions generally). We conclude by making some direct comparisons between the etiological theory and the propensity theory.

On most biological examples, the etiological theory and our propensity theory will yield identical verdicts. There are just two crucial sorts of case on which they part company.

One sort of case which distinguishes the theories is that of the *first* appearance of a character that bestows propensities conducive to survival.

On our theory, the character already has a function, and by bad luck it might not survive, but with luck it may survive, and it may survive *because* it has a function.

On the etiological theory, in contrast, the character does not yet have a function. If it survives, it does not do so because it has a function; but, after time, if it has contributed to survival, the character will have a function.

We think our theory gives a more intuitively comfortable description of such cases, at least in most instances. But there are variants on this theme, on which our theory gives less comfortable results. Suppose a structure exists already and serves no purpose at all. Suppose then that the environment changes, and, as a result, the structure confers a propensity that is conducive to survival. Our theory tells us that we should say that the structure now has a function. Over all, this seems right, but there are cases where it seems counterintuitive. Consider, for instance, the case of heartbeats—that is, the *sound* emitted when the heart beats. In this century, the heartbeat has been used widely to diagnose various ailments; so it has come to be conducive to survival. The propensity theory deems the heartbeat to have the function of alerting doctors. That sounds wrong. The etiological theory says the heartbeat has no such function because it did not evolve for that reason. That sounds plausible.

And yet, we suggest, the reason we are reluctant to grant a function to the heartbeat is not that it lacks an evolutionary past of the required kind. Other characters may lack an evolutionary past, yet may happily invite attribution of a function. Rather, our reluctance to credit the heartbeat with a function stems from the fact that the sound of the heartbeat is an automatic, unavoidable by-product of the pumping action of the heart. And that pumping action serves other purposes. Although the heartbeat does (in some countries, recently) contribute to the survival of the individual, it does not contribute to survival of the character itself. The character—heartbeat—will be present in everyone, whether or not doctors take any notice of it. Although it "contributes" to survival, it is a redundant sort of contribution if it could not fail to be present whether it was making any contribution or not.

Perhaps the propensity theory should be carefully formulated in such a way as to rule out such "automatic" contributions to survival. Nevertheless, we will note only that, although the example of heartbeats seems initially to count against a propensity theory, there are other examples, and wider theoretical considerations, which count in its favor. Further, the case introduces many complications. For these reasons, it cannot be regarded as, in any sense, conclusive.

So much for cases of *new* survival-enhancing characters. There is a dual for these cases: that of characters that were, but are no longer, survival enhancing. These cases, like the former cases, serve to distinguish between the etiological and propensity theories. If a character is no longer survival-enhancing (in the natural habitat), the propensity theory deems it to have no function. The etiological theory, in contrast, deems its function to be whatever it was that it used to be, and was evolved for.

In general, we think the propensity theory gives the better verdict in such cases. Under some formulations, our judgment may be swayed in favor of the etiological theory. We may be inclined to say that the function of a character is to do such and such, but unfortunately this is harmful to the creature these days. Yet surely the crucial fact is, really, that the function *was* to do such and such. It serves no pressing purpose to insist that its function still is to do that. Especially not, once we have passed the threshold over which we redefine the creature's natural habitat. If a character is no longer survival-enhancing, because of a sudden and recent change in environment, we may continue to refer its natural habitat to the past. Consequently, our propensity theory will continue to tie functions to what would be survival-enhancing in the past habitat. In such cases, there will be no conflict between the judgments of our theory and those of the etiological theory.

The test of examples and counterexamples is important. Yet in this case, in the analysis of functions, there is a risk that it will decay into the dull thud of conflicting intuitions. Similarly with intuitions as to how unified should be the analyses of functions for biological characters and for artifacts.

For this reason, we stress the importance of theoretical grounds for preferring the propensity theory. A propensity can play an explanatory causal role, whereas the fact that something has a certain historical origin does not, by itself, play much of an explanatory, causal role. Consequently, the propensity theory has a theoretical advantage, and this gives us a motive for seeking to explain away (or even overrule) apparent counterintuitions.

In a similar way, Darwinian evolutionary theory provides strong theoretical motives for analyzing fitness in a certain way. Our intuitions—our unreflective impulses to make judgments—have a role to play, but not an overriding one.

JOHN BIGELOW
ROBERT PARGETTER

La Trobe University