A Modern History Theory of Functions
Author(s): Peter Godfrey-Smith
Source: *Noûs*, Vol. 28, No. 3 (Sep., 1994), pp. 344-362
Published by: Wiley
Stable URL: https://www.jstor.org/stable/2216063
Accessed: 15-01-2019 20:58 UTC

NOÛS 28:3 (1994) 344–362

A Modern History Theory of Functions

PETER GODFREY-SMITH
Stanford University

I Introduction

Biological functions are dispositions or effects a trait has which explain the recent maintenance of the trait under natural selection. This is the "modern history" approach to functions. The approach is historical because to ascribe a function is to make a claim about the past, but the relevant past is the recent past; modern history rather than ancient.

The modern history view is not new. It is a point upon which much of the functions literature has been converging for the best part of two decades, and there are implicit or partial statements of the view to be found in many writers. This paper aims to make the position entirely explicit, to show how it emerges from the work of other authors, and to claim that it is the right approach to biological functions.

Adopting a modern history position does not solve all the philosophical problems about functions. It deals with a family of questions concerning time and explanation, but there are other difficulties which are quite distinct. The most important of these concern the extent to which functional characterization requires a commitment to some form of adaptationism (Gould and Lewontin 1978). These issues will not be addressed here. Further, as many writers note, "function" is a highly ambiguous term. It is used in a variety of scientific and philosophical theories, several domains of everyday discourse, and there is probably even a plurality of senses current within biology. This paper is concerned with one core biological sense of the term, which is associated with a particular kind of explanation. In this sense a function has some link to an explanation of why the functionally characterized thing exists, in the form it does.

Cummins (1975) argued that functions are properly associated with a different explanatory project, that of explaining how a component in a larger system contributes to the system exhibiting some more complex capacity. Following Millikan (1989b) I suggest that both kinds of functions should be recognized, each associated with a different explanatory project. If it is claimed, for instance, that the function of the myelin sheaths round some brain cells is to make possible efficient long distance conduction of signals, it may not be obvious which explanatory project is involved—that of explaining why the sheath is there, or that of explaining how the brain manages to perform certain tasks. Often the same functions will be assigned by both approaches, but that does not mean the questions are the same.

The aim of this paper is to analyze an existing concept of function, which plays a certain theoretical role in biological science. So the aim is a certain sort of conceptual analysis, a conceptual analysis guided more by the demands imposed by the role the concept of function plays in science, the real weight it bears, than by informal intuitions about the term's application. Also, though I will defend the modern history view within the context of a particular theory of functions which draws on the work of Larry Wright and Ruth Millikan, the overall value of the modern history approach stands independently of many of the details of my theory.[1]

II The Wright Line

Our point of departure is a simple formula proposed by Larry Wright in 1973 and 1976: "The function of X is that particular consequence of its being where it is which explains why it is there" (1976 p.78). That is:

> The function of X is Z iff:
> (i) Z is a consequence (result) of X's being there,
> (ii) X is there because it does (results) in Z. (1976 p.81)

Wright argued that his theory dealt with a broad range of cases, handling both the functions of artifacts and biological entities without significant modification. The function of spider webs is catching prey, because that's the thing they do that explains why they are there; the function of tyre tread is improving traction because that's also the thing it does that explains why its there; and the function of the newspaper under the door is to prevent a draft, for the same reason.

However, Wright's analysis covers more cases than these. Boorse 1976 notes that when a scientist sees a leak in a gas hose, but is rendered unconscious before it can be fixed, on Wright's schema the break has the function of releasing gas. The break is there because it releases gas, keeping the scientist immobilized, and the leaking gas is a consequence of the break in the hose. Similar cases take us even further from the plausible realm of purpose. One might see a small, smooth rock supporting a larger rock in a fast-flowing creek, and note that if it did not hold up that larger rock, it would be washed away, and no longer "be there." But it is not the function of the small rock to support the larger one. The problem here is with the broad range of "X" and "Z," with the need to restrict the kinds of things to which the schema can be applied. A restriction of this kind is a key component of Ruth Millikan's theory (1984, 1989a).

Before moving on however, it is important to recognize Wright's aims. Wright's strategy is to avoid convoluted analysis by trusting many details to pragmatic factors which will apply case by case. For Wright, function hinges directly on explanation, and explanation is pragmatically sensitive in a multitude of ways. There is a sense in which Wright's theory is not an "analysis" of function in the sense that earlier accounts are. Earlier writers were largely concerned with how it can ever be that something's existence can be inferred from its function, given that other things could often have done the same job (Hempel 1965). Without this inference, it was thought there could be no functional explanation. Wright simply insists that with a less demanding, more realistic picture of explanation, it becomes clear that people do explain the presence of things in terms of what they do, and a function is any effect that operates in such an explanation.

Wright also hopes, I suspect, that some natural slack in the notion of function will be mirrored and explained by corresponding slack in the notion of explanation, that the analysis will bend where the concept analyzed naturally bends. Wright's vague formulation of the relevant explanandum—"why its there"—is intended to wrap unsystematically around a variety of explanatory projects, in biology, engineering and everyday life. Nonetheless, counterexamples such as Boorse's do suggest that Wright has backed off too early, and a sensitivity to pragmatics should not prevent us from pushing an analysis as far as we profitably can.

Millikan's analysis, like Wright's, is historical. It locates functions in actual selective histories. The most important sophistication of the historical approach in Millikan 1984 is her detailed treatment of functional *categories*. The first concept she defines is that of a "reproductively established family." A reproductively established family is a group of things generated by a sort of copying. Family members can be copied one off the other, or be common copies off some template, or be generated in the performance of functions by members of another family. These different kinds of copying are all distinguished by Millikan, but the finer divisions are not important here. Call any entities which can be grouped as tokens of a type by these lines of descent by copying, members of a "family." Understand "copy" as a causal matter involving common properties and counterfactuals. The copy is like the copied in certain respects, though it is physically distinct, and if the copied had been different in certain ways, then, as a consequence of causal links from copied *to* copy, the copy would have been different in those ways too (1984 p.20).[2] So two human hearts are members of the same family, as are two frill-necked lizard aggressive displays, two AIDS viruses, and two instances of the acronym "AIDS," assuming that acronym was hit upon only once. But two planets, and two time-slices of a rock or hose are not, as one was not copied off the other, nor are they produced off a common template, and so forth. Functions are only had by family members, and the performance of a function must involve the action of one of the properties copied, one of those properties defining the family.

This restriction deals with many of Boorse's counterexamples, such as the gas hose case. It also removes from the realm of function some cases Wright was concerned to capture, such as the newspaper under the door. However, our project here is to capture the biological usage. Preserving a continuity between biological cases and other domains can be sacrificed.[3]

The next step is to add to this an explanation-schema in the style of Wright. The explanandum is the existence of current members of the family. The explanans is a fact about prior members.

(F1) The function of *m* is to *F* iff:
(i) *m* is a member of family *T*, and
(ii) among the properties copied between members of *T* is property or property cluster *C*, and
(iii) one reason members of *T* such as *m* exist now is the fact that past members of *T* performed *F*, through having *C*.

Most simply, a family member's function is whatever prior members did that explains why current members exist (see also Brandon 1990 p. 188).

It is one of the strengths of the historical approach combined with an appeal to "families" that it can say without strain that some particular thing which is in principle unable to do *F* now, nonetheless has the function to do *F*. It has this function in virtue of its membership in a family which has that function. Whether this member can do *F* is irrelevant to its family membership, as long as it was produced by lines of copying that are generally normal enough. A genetic defect may produce a heart unable to ever pump blood, but if this token was produced in more-or-less the same way as others, it has the function characteristic of the family.

At this point we must confront an issue unrelated to history. It is striking that while analyses such as Wright's and Millikan's permit any activity or power explaining survival to qualify as a function, biologists apparently reserve "function" for activities or powers which are, in some intuitive sense, helpful and constructive. If being inconspicuous and avoiding attention by doing nothing is itself "doing something," then pieces of junk DNA, which sit idly on chromosomes and are never used to direct protein synthesis, have the function to do nothing. That is the thing past tokens of junk DNA types have done, which explains the survival of present tokens. If doing absolutely nothing is a behavior when an animal does it for concealment, why is it not something that junk DNA "does"? Perhaps the function of junk DNA is, alternatively, to be more expensive to get rid of than to retain. But biologists do not describe junk DNA like this; it is the paradigm of something with no function. Similarly, characters which hitchhike genetically on useful traits or persist through developmental inevitability (like male nipples) might, in extended senses, be "doing" things which lead to their survival. So we might consider making some restriction on the selective processes relevant to functional status.

This will not be easy. A simple requirement that the trait do something positive, that the null power is not a power, will not suffice. Beside the cases where biological entities persist through doing nothing, there are positive and selectively salient powers which seem unlikely candidates for functions. As well as junk DNA, which does nothing, there is "selfish DNA" (Orgel and Crick 1980). Selfish DNA can move around within the genome, replicating itself as it goes, and proliferate in a population despite having deleterious effects on individuals carrying it.

Similarly, segregation distorter genes disrupt the special form of cell division (meiosis) which produces eggs and sperm (gametes). Meiosis usually results in a cell with two sets of chromosomes giving rise to four gametes with one set each, and on average a particular type of chromosome will be carried by half the gametes produced. Segregation distorters lever their way into more than their fair half share of gametes, by inducing sperm carrying the rival chromosome to self-destruct as they are formed (Crow 1979). Fruit flies, house mice, grasshoppers, mosquitoes and a variety of plants are known to have segregation distorters in their gene pools. Now, disrupting meiosis is something that segregation distorter genes do, that explains their survival (Lewontin 1962). Further, this explanation appeals to natural selection, at the gametic level; the problem can not be solved by disqualifying traits that survive for non-selective reasons. Disrupting meiosis is not generally claimed to be the genes' *function* though. Should we restrict the powers which can become functions, to exclude these subversive cases?[4]

There are two attitudes we might have to this issue. First, as a question of conceptual analysis, there is not much doubt that biologists typically restrict the powers that can qualify as functions. Many might say we should then change the selective theory of functions to include this factor. An obvious move is to bring in some reference to the goals of some larger system. Disrupting meiosis makes no contribution to the goals of individuals bearing segregation distorter genes, so this is not a function.

An appeal to goals is certainly a step backwards however. So we might consider a more aggressive attitude to the problem. It may be that many biologists reserve "function" for powers with some intuitively benign nature, and withhold it from more subversive activities, with there being no theoretically principled reason for this distinction. Some hold that biology since the 1960's has produced, for better or worse, an increasingly cynical view of the coalitions that make up organisms (Dawkins 1982, Buss 1987), families (Trivers 1974), and larger groups (Williams 1966, Hamilton 1971). The feeling that functions must involve harmonious interactions may, from this point of view, be a holdover from an earlier, more truly teleological view of nature. It might be claimed that the theoretically important category of properties, the category our concept of function should be tailored to, is simply the category of selectively salient powers and dispositions.[5] If so, we should remain with the simpler analysis that allows any survival-enhancing power, however subversive, to qualify as a function.

Although some may favor this more heartless approach I will adopt a third, intermediate position. Consider first another counterintuitive consequence of an unembellished selective account: whole organisms, like people, have functions. Past tokens of people did things—survived and reproduced—that explain why current tokens are here. Hence, we have the function to survive and reproduce. This usage seems odd—note that these are not functions people might have with respect to some social group, they are functions people just have, individually. One way to exclude both people as bearers of functions and also exclude disruption of meiosis as a function of segregation distorters is to stipulate that (i) the functionally characterized structure must reside within a larger biologically real system, and (ii) the explanation of the selection of the functionally characterized structure must go via a positive contribution to the fitness of the larger system. My account here resembles that of Brandon, who requires that a functional trait increase the "relative adaptedness of [its] possessors" (1990 p.188). Brandon requires not just selective salience, but selective salience which goes via the fitness of a larger system "possessing" the trait.

The catalog of "real systems" is taken from biology, and clarifying the catalog is part of the units of selection problem.[6] Individuals, kin groups and perhaps populations and species might be examples of these systems. Thus hearts reside within people, and survive by aiding people's fitness. But people, considered individually, reside within no such systems. There may, however, be groups within which people do things which contribute to the selection of the group, and then people would have functions.

Similarly, segregation distorter genes do not have the function of disrupting meiosis, because their proliferation under selection does not occur through a positive contribution to the fitness of individuals bearing these genes. Indeed, many segregation distorters, when present in two copies, greatly impede the fitness of their carriers. On the other hand, as some readers may have felt earlier, there could well be functional characterization of *parts* of segregation distorter genes or gene combinations. Some part of the gene or combination might have its current presence explained by the fact that it has been selected for carrying out some part of the segregation distortion project. Crow (1979) distinguishes two genes which cooperate to produce segregation distortion in fruit flies. The "*S*" gene produces sabotage in sperm, and the "*R*" gene stops the chromosome that the *S* and *R* are on from sabotaging itself. So a chromosome with *S* but no *R* sabotages itself, and a chromosome with *R* but no *S* does not distort, but is immune to distortion by its rival. Here the segregation distorting chromosome is the larger system, and the selective explanation of *S* goes via the explanation of the success of the whole chromosome. *S* has the function of sperm sabotage, and it has this function with reference to the segregation distortion gene complex. The selection of *R* is only partly an explanation in terms of the selection of the distorter chromosome, as *R* is useful without *S*, once the population contains some chromosomes with *S*. So *R* has the function of preventing sabotage, and it has this function with reference to two larger units, the segregation distorter complex and the individual.

It is important that not all failures on the part of evolution to produce intuitively well-engineered animals disqualify selective episodes from bestowing functions. A question sometimes arises concerning the status of traits which are explained in terms of some forms of sexual selection. If it is true that sexual selection can operate through females favoring characteristics in males which have no other benefit or use to the male (Fisher 1930, Lande 1981), then the explanation of a bird's long tail is not an explanation in terms of anything intuitively useful the tail does. The explanation is simply that females prefer long tails (Andersson 1982). Once a female preference gets established, for any reason, it can be sustained and made stronger through the association of the gene for the preference in females (unexpressed in males) and the gene for the preferred trait (unexpressed in females). The preference leads to the selection of long tails, and the selection of long tails leads to the strengthening of the associated preference. The long tail could be a hindrance elsewhere in life. Consequently, some biologists hesitate to describe the tail as an adaptation, and functional in the ordinary sense: "Runaway sexual selection is a fascinating example of how selection may proceed without adaptation" (Futayama 1986 p.278). On the present account however the tail has the function to attract females. It has been selected because of that power, and this explanation goes via the augmentation of the individual's fitness.[7]

Here is an amended definition:

(F2) The function of *m* is to *F* iff:
  (i) *m* is a member of family *T*,
  (ii) members of family *T* are components of biologically real systems of type *S*,
  (iii) among the properties copied between members of *T* is property or property cluster *C*,
  (iv) one reason members of *T* such as *m* exist now is the fact that past members of *T* were successful under selection, through positively contributing to the fitness of systems of type *S*, and
  (v) members of *T* were selected because they did *F*, through having *C*.

III Looking Forward

Although philosophers have discussed a variety of intuitive problems with the view that functions derive from a selective history (Boorse 1976), the most damaging charge against this view derives from the biological literature, from the wide acceptance of the distinctions made in "Tinbergen's Four Questions."

It is common in ethology and behavioral ecology to distinguish four questions "why?" we can ask about a behavior. Someone who asks why frill-necked lizards extend the skin around their necks so spectacularly might want an answer:

1. In terms of the physiological *mechanisms* and the physical stimuli that lead to the behavior.
2. In terms of the current *functions* of the behavior.
3. In terms of the evolutionary *history* of the behavior.
4. In terms of the *development* of the behavior in the life of the individual lizard.

This four-way distinction is usually attributed to Tinbergen 1963. Tinbergen in turn credits Julian Huxley with distinguishing questions 1–3, and adds question 4. Tinbergen, it must be admitted, uses the term "survival value" rather than "function" in the official formulation of question 2. But generally he uses these two expressions interchangeably (1963 p.417, 420).

Tinbergen's distinctions are often endorsed in the opening pages of books about animal behavior (Krebs and Davies 1987 p.5, Halliday and Slater 1983 p.vii, and see also Horan 1989). This is clearly an embarrassment for any historical theory of function which seeks to capture biological usage: on the historical view there should be three questions, not four, as the functional question *is* a question about evolutionary history, as long as the rest of (F2) above is satisfied. Related distinctions with this separation between function and history are found elsewhere in evolutionary writings as well. Mayr 1961 distinguishes "functional" from "evolutionary" biology, and Futuyma's widely used textbook echoes Mayr in dividing the study of biology into functional and historical "modes" (1986 p.286).[8]

There are various ways to respond to this problem. Many ahistorical usages of "function" are probably best understood as referring to Cummins' functions. However, it is common for writers to both regard functions as ahistorical *and* regard them as intrinsically tied to natural selection, sometimes via the expression "survival value." This supports the proposal of a number of writers that functions involve not actual selective histories, but probable futures of selective success, or atemporal dispositions to succeed. Tinbergen may have accepted such a view: "the student of survival value, so-to-speak, looks 'forward in time'" (1963 p.418). Tinbergen (p.428) also casts the question about a structure's function as a question about how deviations from the actual structure *would* lower the fitness of the bearer. John Staddon concurs (1987 p.195). One way to develop this approach is with an appeal to propensities.

Bigelow and Pargetter (1987) develop a theory of functions modelled explicitly on the widely accepted propensity view of fitness (Mills and Beatty 1979).

The propensity view of fitness claims that the fitness of an individual is not the actual fact of its reproductive success, but its propensity to have a certain degree of reproductive success. Similarly, Bigelow and Pargetter claim, functions should be understood as dispositions or propensities to succeed under natural selection. "Something has a (biological) function just when it confers a survival-enhancing propensity on a creature that possesses it" (1987 p.192).

The propensity view is not satisfactory, though its failure performs the valuable service of narrowing the discussion down, along with Tinbergen's Four Questions, to a point where the modern history view will become compelling. I will discuss first some internal difficulties with the propensity view and then argue that the whole forward-looking approach is on the wrong track.[9]

The central internal problem is that as one tries to fill in some more details, the theory tends to go in one or other of three undesirable directions. It can become enmeshed in strong counterfactual commitments. Alternately, it draws on the historical facts it sought to avoid. Or thirdly it makes the wrong kinds of demands on the future. Putting it briefly: propensities to be selected and survive bestow functions, but, the questions swarm: survive where? be selected over *what*? Bigelow and Pargetter address the first question, admitting that their account "must be relativized to an environment" (p.192). The context assumed is the creature's "natural habitat." "Natural habitat," it appears, is understood historically by Bigelow and Pargetter. The statistically most common context for a trait now might be odd and unnatural (Neander 1991b).

More worrying is the question of the competitors that have a propensity to be ousted from the population by the trait we are interested in. Bigelow and Pargetter make no mention of the fact that claims about propensities to do well under natural selection are surely always comparative claims. A trait does not have a propensity to be selected and survive simpliciter, but always a propensity to be selected over some range of alternatives. Evolution is driven by differences in *relative* fitness. Bigelow and Pargetter cannot claim that current useful traits would triumph over any possible alternatives. Which are the relevant ones? Those alternatives genetically attainable (given mutation rates, population structure, other constraints...) now? Those that could enter the fray during the next thousand years? Those that could enter the fray if the ozone layer goes and mutation rates are elevated? If Bigelow and Pargetter think there is a range of alternatives, and circumstances of selection, appropriate to the trait in question independently of history, they are making strong modal commitments. These might be avoided with an appeal to what is most likely to happen in the actual future, but then problems are created by (what appear to be) irrelevant contingent features of this future. If a trait is adaptive, but doomed because of linkage to something bad, then it is not likely to survive. But this should not make a trait itself non-functional.

So, though the propensity theory is tailored to avoid dragging up the past, the propensities involved must either make tacit reference to millennia gone by, inappropriate predictions about the future, or questionable modal commitments about relevant ranges of alternatives and circumstances of selection. These internal problems are important, because it is easy to think that propensity views are somehow more economical than analyses appealing to the past. Still, the propensity view has recommendations. It does seem to be a way to accommodate the intuition that functions derive from selection with the observation that many biologists keep functional and historical questions separated. In addition, I am often told that no matter how questionable philosophers may find the modal commitments outlined above, many biologists constantly talk as if these facts are quite unproblematic and accessible. It is difficult to work out the right attitude to such a datum. Further, one principled way to deal with these internal problems is to fashion a mixed theory, using the basic propensity format with an appeal to history to answer the objections raised above. (This mixing was suggested to me by Elisabeth Lloyd).

The mixed theory claims that functions derive from propensities to be selected, but all the factors that Bigelow and Pargetter left vague are understood historically. The relevant ecological conditions are the actual ones that obtained during the development of the trait. The range of alternatives the trait has a propensity to be selected over are the ones it actually triumphed over, and continues to be selected over. The propensity that bestows functions is strictly atemporal; a trait is held to have a certain advantage under certain conditions over certain rivals. But these conditions and rivals are determined by the actual world. So it does seem likely that the propensity approach can be developed in a coherent way, at the price of narrowing the gap between it and the historical view. This is the general form of the contemporary functions debate: each theory is made more plausible by setting it on a course of convergence with its rivals.

There is, however, a more important problem with propensity theories, and other forward-looking views. These theories inevitably distort our understanding of functional explanation. In the first section I claimed that the sense of function under discussion is a sense linked in some way to explanations of why the functionally characterized entity exists, or exists in the form it does. The most straightforward way to envisage this link, which I have been assuming, is to say that functions are used in explanations of why the functionally characterized thing exists now. If this is granted, and the explanation is understood causally, then there is a simple argument against propensity views. The only events that can explain why a trait is around now are events in the past. Forward-looking accounts claim that functions are not bestowed by facts about the past, but rather by how things are in the present. But then appealing to a function cannot itself explain the fact that the trait exists now. If the environment is uniform, then present propensities to do well under selection may be a good guide to actual prior episodes of selection. But this epistemological point does not alter the fact that it is not the present propensities, but the prior episodes, that are causally responsible for how things are now (see also Millikan 1989b, Neander 1991a).[10]

I do not claim that Bigelow and Pargetter have missed this straightforward point. On their view, there is a problem with the background assumptions I have made about the explanatory role of functions, and which the argument above assumes. Bigelow and Pargetter claim that if the fact that some effect is a function itself depends on the fact that this effect explains the survival of the trait in question, if the assignment of a function is always retrospective in this way, "then 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" (1987 p.190). This vacuity problem can be solved, according to Bigelow and Pargetter, if functions are understood as propensities. These propensities can be used to explain the existence of a trait in the present if we claim, in addition, that the propensities in question did exist in the past, and were causally active in the past. This postulation of the past action of the propensities is an extra claim; it is not guaranteed by the mere fact that the effects in question are functions.

Bigelow and Pargetter's claims about explanatory vacuity and the historical view have been criticized effectively by Sandra Mitchell (1993) She points out that if we say "Trait X persisted because it had a consequence responsible for its selection and consequent evolution," this is only vacuous if we read "persisted" as meaning "evolved by natural selection." That is, it is only vacuous if we assume that the *only* mechanism which could explain some trait being around today is natural selection, though in fact there are alternative evolutionary forces which could play this explanatory role (1993 p.253–54). This is correct, and it shows that Bigelow and Pargetter's argument about the vacuity of the historical view assumes an implausible adaptationism. There is also another objection to Bigelow and Pargetter's claim, which is compatible with even the strongest adaptationism. On the historical view and with the assumption of adaptationism, it will be truly vacuous to say that *X* persisted because it serves *some* function, because we are assuming that this is the only possible type of explanation. But even against this background it will of course not be vacuous to say that *X* persisted because it provided effective camouflage, or because it attracted mates, or because it conserved heat. Neither is it vacuous to say that the trait persisted because some specific effect was its *function*. If the historical theorist says "*X* persisted because its function was to conserve heat," this is to be translated into something which is ungainly, and contains a redundancy—"*X* persisted because its actually-selected effect was that of conserving heat." But this is not vacuous; it does contain a real explanation, though to express it this way mentions the explanatoriness of the effect twice. So this is not the most natural mode of expression for the historical view; on that view the sentence "The function of *X* is to conserve heat" is itself explanatory, and if someone is asked "Why is *X* there?" they can reply by simply citing the function. This is not possible at all on the propensity view. On the propensity view, a functional explanation must give a function and also make an additional claim that the function was causally active in the past.

So despite what Bigelow and Pargetter claim, as long as "a given function" is understood to refer to some specific task or benefit, it is not trivial to say that "the character has persisted because it serves a given function," even assuming adaptationism. This, along with Mitchell's argument, shows that there is no vacuity problem with the background assumptions about explanation that proponents of the historical view make. It is possible to retain the explanatory force of function ascriptions, along with the philosophically attractive view, argued by Wright, that actual explanatory salience is exactly what *distinguishes* functions from mere effects.

A "forward-looking" approach to functions has also been endorsed by Barbara Horan (1989), but the claims she makes about explanation are more problematic than those of Bigelow and Pargetter. Horan says "questions about the function of a given pattern of social behavior are a way of asking how that behavior enhances the fitness of an individual who engages in it" (1989 p.135). Nevertheless, she claims soon after that the presence of a trait like a social behavior can be explained by an attribution of a function to that behavior. The model of explanation she applies, citing G.A. Cohen, is called a "consequence explanation." Consequence explanations use laws of the form: "If (if *C* then *E*), then *C*." In the present context: "if a behavior pattern would increase individual fitness, individuals will come to display that behavior" (1989 p.136).

This is trying to have it both ways. It is true that useful things a behavior does now can lead to its prevalence in the future. So forward-looking functions may predict and explain the future prevalence of a trait. But if the explanandum is how things are now, nothing present or future can be the explanans. Only the past will do. Of course, traits that are useful now were often useful then, so we can often infer that a propensity existing now was also causally active then. But if so, it is explanatory with respect to the present *because* it was causally active then. To claim that present usefulness in itself explains the morphologies and behaviors organisms presently display, and to build this into an account of functions, is to distort the explanatory structure of evolutionary theory.

IV The Modern History Theory

It might appear that we are painting ourselves into an analytical corner. Historical analyses are unacceptable because they fail to respect an apparently important distinction in biology between functional and evolutionary explanation. Forward-looking analyses are unacceptable because they distort our understanding of functions' explanatory role. In fact there are several options available at this point. Bechtel (1989) suggests that we retain a forward-looking account of functions while giving up our prior conception of functional explanation. We might, alternately, claim that functional explanation just is evolutionary explanation, and banish other notions of function (except for Cummins') as creatures of teleological darkness. A third option is to analyze functional explanation as a particular *kind* of evolutionary explanation. One alternative here is to regard a functional explanation as a selective explanation which satisfies (F2) above, hence a subset of evolutionary explanation. The option I prefer, however, is to construe functional explanation more narrowly still.

This brings us, at last, to the modern history view: functions are dispositions and powers which explain the recent maintenance of a trait in a selective context. Several people have already said, in effect, that this is the answer, but these people either make the suggestion in passing (Kitcher 1990), or more often, they only say it some of the time. Horan says "to explain the maintenance of a trait in a species, one gives a functional explanation" (1989 p.135), but insists on an atemporal construal of this explanation. And consider this remark of Millikan's, in response to Horan:

> If natural selection accounts for a trait, that is something that happened in the past, but the past may have been, as it were, "only yesterday." Indeed, *usually* the relevant past is only yesterday: the *main business* of natural selection is steady maintenance of useful traits against new intruders in the gene pool. But only yesterday is not outside of time. (1989b p.173)

We need not endorse the claim about the "main business" of natural selection; whether or not maintaining traits is the main business of selection, it is one important kind of selection. It might be important enough to make this a constitutive part of the concept of function. Millikan does not take this step; her historical account does not *build into* functions the historically recent nature of the relevant selective episodes. Indeed, in her 1984 treatment she explicitly allows powers which were important in ancient history, but not in modern history, to be functions (1984 p.32). In the 1989b treatment her emphasis is different, and she claims the relevant past is "*usually*" only yesterday. But perhaps, as far as functions go, it *must* be only yesterday.

The modern history view does not respect the letter of Tinbergen's Four Questions, but it is faithful to their spirit. Tinbergen makes the modern/ancient history distinction himself (1963 pp.428–29), but he regards *both* these explanations as "evolutionary" rather than functional. This puts two distinct questions under one head, however, as well as leaving the explanatory significance of functions in the dark. From the present viewpoint, the "evolutionary" question is the question about the forces which originally built the structure or trait in question. This may or may not be a selective explanation, and this explanation might be different from the explanation of why the trait has recently been maintained in the population.

Some might wonder how recent the selective episodes relevant to functional status have to be. The answer is not in terms of a fixed time—a week, or a thousand years. Relevance fades. Episodes of selection become increasingly irrelevant to an assignment of functions at some time, the further away we get. The modern history view does, we must recognize, involve substantial biological commitments. Perhaps traits are, as a matter of biological fact, retained largely through various kinds of inertia. Perhaps there is not constant phenotypic variation in many characters, or new variants are eliminated primarily for non-selective reasons. That is, perhaps many traits around now are not around because of things they have been doing. Then many modern-historical function statements will be false. If functions are to be understood as explanatory, in Wright's sense, there is no avoiding risks of this sort.

One way to support the modern history view of function is to demonstrate that the category of explanation it distinguishes is a theoretically principled one. This can be done by focusing on traits for which the modern historical explanation and the ancient historical explanation diverge, so the selective forces salient in the origin of the trait are different from those salient in the recent maintenance of the trait. Here is where a distinctively functional style of characterization—in the modern history sense—can be seen to be useful.

The importance of the distinction between modern and ancient evolutionary explanations is discussed, in support of an analysis of function quite opposed to mine, in Gould and Vrba 1982. The central concern of Gould and Vrba is a distinction between adaptations and "exaptations" (their coinage). They understand adaptations as characters shaped by natural selection for the role they perform now. Exaptations are characters built originally by selection for one job, or characters with no direct selective explanation at all, which have since been coopted for a new use. This analysis has consequences for their concept of function; only adaptations have functions, and exaptations have "effects." Gould and Vrba do not discuss the recent past, as distinct from the present, so I am uncertain how they would classify modern-historical functions. Generally they seem to understand effects-of-exaptations as propensities (1982 p.6). Their effects-of-exaptations correspond to the functions of Bigelow and Pargetter, and Horan. It should be clear why I think their way of dividing the cases is inadequate: modern history and ancient history can *both* furnish genuine explanations, which we should distinguish, for why something exists now, while present propensities cannot themselves furnish such explanations.

Gould and Vrba's central point is the importance of cases where a trait's original and current uses diverge. But these are also cases where the selective forces that built a character and those maintaining it in the recent past diverge, so they also illustrate the origin/function distinction as I understand it. Gould and Vrba make two claims about such cases. Firstly, there are many of them, and secondly, the cases are theoretically significant. The co-opting of existing traits for new uses is important in the development of complex and novel adaptive characters.

Feathers, it has been argued, did not originate as adaptations for flight. The earliest known bird *Archaeopteryx* did not have the skeleton for anything beyond very rudimentary flight, but was well-covered with feathers. It has been claimed that feathers originated as insulation, and only later were coopted for flight

(Gould and Vrba 1982 p.7 cite Ostrom 1979). Thus the question about the evolutionary origin of feathers is answered in terms of selection for effective insulation, but if we ask today about the function of feathers, in a sub-tropical bird for instance, the answer appeals to the reason feathers have recently been maintained—their facilitating flight.

A similar story can be told about the development of bone. Bone is essential as a support for land-dwelling vertebrates, but it developed in sea animals well before it could be put to its modern use. Gould and Vrba discuss the hypothesis that bone was developed as store of phosphates needed for metabolic activity (Halstead 1969). In this case, the original use continues, and bone functions in modern vertebrates as storage for mineral ions, including phosphate ions, as well as support.

Gould and Vrba's examples can be augmented easily. The electric eel's ability to kill prey and defend itself with electric shocks is a development of the weaker electric abilities of other fish, which generate electric fields as part of a perceptual system, used in orientation and communication (Futayama 1986 pp.423–24). Shepherd (1988 p.67) discusses a suggestion made by J.B.S. Haldane about the origin of neurotransmitters, the chemicals whose function now is passing signals between neurons in the brain. Haldane suggested that these chemicals may have developed originally as chemical messengers between individuals. There are a number of neurotransmitters which can induce effects on other organisms.

A final illustration of the importance of the distinction between originating and maintaining selection is found in some of the literature applying game theory to animal behavior (Maynard Smith 1982).[11] An ESS, or evolutionarily stable strategy, is a strategy which, once prevalent in a population, cannot be invaded by rival strategies. However, an ESS need not be a strategy that can evolve from scratch in any situation. Often a critical mass of like-minded individuals is needed before a strategy becomes stable. Thus to explain a behavior by showing it to be an ESS is not necessarily to explain how that behavior originally became established. Rather, it is to point to the selective pressures responsible for the recent maintenance of the strategy in the population.[12]

The point is not just the apparent commonality of a divergence between modern and ancient history, but the fact that this distinction has sufficient theoretical importance to justify its place in an analysis of functions.

One final problem must be discussed, which can be introduced with a feature of Wright's analysis. It is initially perplexing that Wright uses the present tense in the expression: "*X* is there because it does (results in) *Z*" (1976 p.81). If his account is historical ("etiological"), why does he not make it explicit that the performances of *Z* that explain the presence of *X*'s are in the past?

> In general, when we explain something by appeal to a causal principle, the tense of the operative verb is determined by whether or not the principle still holds at the time

> the explanation is given… . We might say, for example, "The Titanic sank because when you tear a hole that size in the bow of a ship it sinks," using the verb "to sink" in the present tense even though the sinking in question took place in the past… . If we were to throw the statement into the past tense it would imply that nowadays one could get away with tearing a hole that size in the bow of a ship without it sinking. (1976 pp.89–90)

Wright requires that the effects appealed to in a functional explanation still exist at the time of the functional ascription, and these effects must still have the same causal efficacy that they have had in the past. If this means that the structure in question must now have a propensity to continue to be selected for the same reason that it was selected in the past, Wright's account converges with that of Gould and Vrba, who demand that functions presently "promote fitness" (1982 p.6).

Should the modern history view include these extra requirements? In my view, there may be good reason to require that the trait still be able to do now what it was selected for doing, but we should not require that the trait also have the same propensity to succeed under selection that it has had in the past. This problem is less pressing for the modern history view than for other historical views. If a trait has very recently been selected for doing *F*, it will tend to still be able to do *F* now. As it is *possible* for it (the type) to be unable to do *F* now, no matter how recently it has been selected for doing *F*, it is probably reasonable to add an extra clause requiring the continuation of the disposition into the strict present.[13] Whichever way one goes here, it is an advantage of the modern history view that these uncooperative cases should be made very rare.

Here is my final attempt at a definition of function.

(F3) The function of *m* is to *F* iff:
  (i) *m* is a member of family *T*,
  (ii) members of family *T* are components of biologically real systems of type *S*,
  (iii) among the properties copied between members of *T* is property or property cluster *C*, which can do *F*,
  (iv) one reason members of *T* such as *m* exist now is the fact that past members of *T* were successful under selection in the recent past, through positively contributing to the fitness of systems of type *S*, and
  (v) members of *T* were selected because they did *F*, through having *C*.

Much of this definition is proposed tentatively. The most important part is the appeal to modern history, which can also be incorporated in other theories of functions. The central recommendation of the modern history view is the fact that it accounts for the explanatory force of function ascriptions, but does this while making sense of the biological distinction between "functional" and "historical" explanation. It is a theory which steers a principled middle course.

Notes

Acknowledgment: This work developed largely out of a series of discussions with Phillip Kitcher. Along with many of the ideas, the term "modern history theory" was his, though he should not be taken to endorse (or reject) the modern history view. I have also benefitted from discussions with Elisabeth Lloyd, Ruth Millikan, Sandra Mitchell and everyone at Kathleen Akins' Functions Reading Group. An anonymous referee for *Noûs* made a number of valuable criticisms of earlier drafts. I would also like to thank the University of Sydney for generous financial support during the period when most of this work was done.

[1] Neander (1991a), Mitchell (1989, 1993) and Brandon (1990) have defended theories of functions running along similar lines. Sober's (1984) analysis of adaptation is also a relative.

[2] Those familiar with some units of selection debates in philosophy of biology will note that family members need not be replicators: see Dawkins 1982, Hull 1981.

My definition of copying is not supposed to be airtight, and may be too inclusive. Kim Sterelny suggested that it lets in molecular structures in a crystal lattice, for instance, though it is not so certain that this case should be kept out. See Millikan 1984 for more details.

[3] Millikan presents her 1984 account as a stipulative definition, not an analysis of an existing concept, so this is not a problem for her. It is also important that Millikan's restrictions do not prevent the analysis being applied to artifacts generated by copying in the right ways.

[4] The treatment in Millikan 1984 fudges here. Millikan's official definition of function begins with a stipulated function F, and explains why something has this function F. Can *any* activity or power qualify as function F, as long as it promotes survival? If not, Millikan owes us an account of what sorts of properties can be functions. If on the other hand she allows any power to be a function, then why does she take the indirect route, of starting with a function to be fulfilled and then explaining why one structure, rather than a rival, has this function as its own?

[5] Most philosophical commentators on an earlier draft of this material inclined towards the heartless line on this question.

[6] In the terms of the units of selection debate, the larger system needs to be a real *interactor* (Hull 1981, Lloyd 1988, Brandon 1990).

[7] Wright 1976 discusses the possibility of an appeal to the broader system (p.106). He dismisses it firmly (though this fails to prevent other writers from attributing such an appeal to him: Nagel 1977 p.283, Hampe and Morgan 1988 p.123). Wright however does not discuss examples like those causing trouble in the present discussion.

[8] A puzzling case is George Williams (1966) Williams is usually regarded as an advocate of a Wright-style account of functions (Boorse 1976 p.85, Wright 1976 pp.92–93), as suggested by this well-known passage: "One should never imply that an effect is a function unless he can show that it is produced by design [natural selection] and not by happenstance" (1966 p.261). But when Williams lays down principles for the general study of adaptation, he seems to imply that the basic fact of something's having a function is not a historical fact. It appears that the "prime" question asked about a character in such a study—"What is its function?"—is answered in terms of contributions to goals (1966 p.258, citing Pittendrigh 1958). The *second* question asked is the historical one about selection (p.259, see also p.264). Williams does goes on to say that an activity is not a function unless it was produced by design rather than chance (p.261, quoted above). So the ahistorical nature of the "prime" question might be merely epistemological.

[9] The version of the propensity view I am discussing is based on the survival propensities of *character types*. The propensity is possessed by human hearts as a type, not by individual hearts, and not by individual people. Bigelow and Pargetter are not consistent here. Sometimes they talk about the survival of the individual bearing the functionally characterized trait (1987 p.192). But later, when speaking more strictly, they focus explicitly on the character type (p.195, see also p.194). On my reading, their talk of the "survival" of individuals is really talk of individuals' inclusive fitness (in the biological cases at least). Sandra Mitchell pointed out to me that if their propensities are read as belonging to individual trait-bearers, their theory is more like a classical goal theory. Admittedly, they do regard their account as a "cousin" of goal theories (p.182). Neither interpretation squares with everything they say, but this exegetical question is less important than the theoretical issue of the viability of a propensity-based selective account.

[10] Focusing on causal explanation in this way also makes it clear why the selective advantage relevant to functional status cannot be understood with reference to a range of counterfactual alterna-

tive traits, as opposed to actual ones, as some propensity views might maintain. Only competition with actual, past rivals is causally relevant in explaining why a trait exists today.

[11] I am indebted to Philip Kitcher for this point.

[12] The distinction between the original establishment and the maintenance of a strategy is stressed, for instance, in Axelrod and Hamilton's well-known discussion of the properties of tit-for-tat in the iterated prisoner's dilemma (1981).

[13] This suggestion is made cautiously—perhaps all these additional requirements are ill-advised (Neander 1991b p.183).

References

Andersson, M. (1982) Female Choice Selects for Extreme Tail Length in a Widowbird. *Nature* 299: 818–20.

Axelrod, R. and W. Hamilton (1981) The Evolution of Cooperation. *Science* 211: 1390–6.

Bechtel, W. (1989) Functional Analyses and their Justification. *Biology and Philosophy* 4: 159–162.

Bigelow, J. and R. Pargetter (1987) Functions. *Journal of Philosophy* 84: 181–197.

Boorse, C. (1976) Wright on Functions. *Philosophical Review* 85: 70–86.

Brandon, R. (1990) *Adaptation and Environment*. Princeton: Princeton University Press.

Brandon, R. and R. Burian., eds. (1984) *Genes, Organisms, Populations: Controversies over the Units of Selection*. Cambridge, MA: MIT Press.

Buss, L. (1987) *The Evolution of Individuality*. Princeton: Princeton University Press.

Crow, J. (1979) Genes that Violate Mendel's Rules. *Scientific American* 240: 134–146.

Cummins, R. (1975) Functional Analysis. *Journal of Philosophy* 72: 741–765.

Dawkins, R. (1982) *The Extended Phenotype*. Oxford: Oxford University Press.

Fisher, R. A. (1930) *The Genetical Theory of Natural Selection*. Oxford: Clarendon.

Futayama, D. (1986) *Evolutionary Biology*. 2nd edition. Sunderland: Sinauer.

Gould, S. J. and R. C. Lewontin (1978) The Spandrells of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Program. *Proceedings of the Royal Society, London* 205: 581–598.

Gould, S. J. and E. Vrba. (1982) Exaptation—a Missing Term in the Science of Form. *Paleobiology* 8: 4–15.

Halliday, T. R. and P. J. B. Slater, eds. (1983) *Animal Behavior, Vol.2: Communication*. New York: Freeman.

Halstead, L. B. (1969) *The Pattern of Vertebrate Evolution*. Edinburgh: Oliver and Boyd.

Hamilton, W. D. (1971) Geometry for the Selfish Herd. *Journal of Theoretical Biology* 31: 295–311.

Hampe, M. and S. R. Morgan. (1988) Two Consequences of Richard Dawkins' View of Genes and Organisms. *Studies in the History and Philosophy of Science* 19: 119–138.

Hempel, C. G. (1965) The Logic of Functional Analysis. In *Aspects of Scientific Explanation*. New York: Free Press.

Horan, B. (1989) Functional Explanations in Sociobiology. *Biology and Philosophy* 4: 131–158.

Hull, D. (1981) Units of Evolution: a Metaphysical Essay. Reprinted in Brandon and Burian 1984.

Kitcher, P. S. (1990) Developmental Decomposition and the Future of Human Behavioral Ecology. *Philosophy of Science* 57: 96–117.

Krebs J. and N. Davies (1987) *An Introduction to Behavioural Ecology*, 2nd edition. Oxford: Blackwell.

Lewontin, R. C. (1962) Interdeme Selection Controlling a Polymorphism in the House Mouse. *American Naturalist* 96: 65–78.

Lloyd, E. A. (1988) *The Structure and Confirmation of Evolutionary Theory*. New York: Greenwood Press.

Maynard Smith, J . (1982) *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Mayr, E. (1961) Cause and effect in biology. *Science* 134: 1501–1506.

Millikan, R. G. (1984) *Language, Thought, and Other Biological Categories*. Cambridge, MA.: MIT Press.

Millikan, R. G. (1989a) In Defence of Proper Functions. *Philosophy of Science* 56: 288–302.

Millikan, R. G. (1989b) An Ambiguity in the Notion "Function." *Biology and Philosophy* 4: 172–176.

Mills, S. and J. Beatty (1979) The Propensity Interpretation of Fitness. *Philosophy of Science* 46: 263–286.

Mitchell, S. (1989) The Causal Background of Functional Explanation. *International Studies in the Philosophy of Science*. 3: 213–229.

Mitchell, S. (1993) Dispositions or Etiologies? A Comment on Bigelow and Pargetter. *Journal of Philosophy* 90: 249–59.

Nagel, E. (1977) Teleology Revisited: The Dewey Lectures 1977. (1) Goal-directed Processes in Biology. (2) Functional Explanations in Biology. *Journal of Philosophy* 74: 261–301.

Neander, K. (1991a) The Teleological Notion of "Function." *Australasian Journal of Philosophy* 69: 454–468.

Neander, K. (1991b) Functions as Selected Effects: The Conceptual Analyst's Defence. *Philosophy of Science* 58: 168–184.

Orgel, L. E. and F. H. C. Crick (1980) Selfish DNA; the Ultimate Parasite. *Nature* 284: 604–606.

Ostrom, J. H. (1979) Bird flight: How Did it Begin? *American Scientist* 67: 46–56.

Pittendrigh, C. S. (1958) Adaptation, Natural Selection, and Behavior. In A. Roe and G. G. Simpson, eds, *Behavior and Evolution*. New Haven: Yale University Press.

Shepherd, G. M. (1988) *Neurobiology*, 2nd edition. Oxford: Oxford University Press.

Sober, E. (1984) *The Nature of Selection*. Cambridge, MA: MIT Press.

Staddon, J. E. R. (1987) Optimality Theory and Behavior. In J. Dupre, ed. *The Latest on the Best: Essays on Evolution and Optimality*. Cambridge, MA: MIT Press.

Tinbergen, N. (1963) On the Aims and Methods of Ethology. *Zeitschrift für Tierpsychologie*. 20: 410–33

Trivers, R. (1974) Parent-offspring Conflict. *American Zoologist* 14: 249–64.

Williams, G. C. (1966) *Adaptation and Natural Selection*. Princeton: Princeton University Press.

Wright, L. (1973) Functions. *Philosophical Review* 82: 139–168.

Wright, L. (1976) *Teleological Explanations*. Berkeley: University of California Press.