Chapter 6: The survival value approach

6.1 Introduction

In this chapter I am concerned with the survival value approach to function attributions. I argue that this approach is on the right track as an analysis of the meaning of the notion of 'survival value' (function$_3$), but that the current proposals fail to account for the explanatory use of appeals to survival value. In addition, I elaborate my own account of the notion of 'survival value' (which I exposed in section 2.2.3) and defend it against criticisms raised against the survival value approach. My account of the explanatory use of attributions of survival value (function$_3$ attributions) is postponed to chapter 8.

According to the survival value approach a function is an effect of an item, behaviour or trait that contributes to the survival, reproduction or fitness of the organisms that have that item, behaviour or trait. For example, on this account propagating the blood is a function of the heart because propagating the blood contributes to the survival of the organisms that have hearts. This contribution is typically explicated in terms of a counterfactual conditional: an effect has survival value if the organisms that produce it would survive or reproduce less well or would have a lower fitness if they would not produce that effect. If the heart did not propagate the blood, the organisms that have it would survive less well. Survival value accounts are offered by John Canfield (1964, 1965), William Wimsatt (1972), Michael Ruse (1973), John Bigelow and Robert Pargetter (1987), and Barbara Horan (1989). Important criticisms of one or more variants of the survival value account are presented by Harry Frankfurt and Brian Poole (1966), Lowell Nissen (1970), Kenneth Baublys (1975), Ruth Millikan (1989a, 1993a), and Karen Neander (1991a,b).

As I will show, both the adherents and the critics of a survival value approach tend to confuse attributions of causal roles (function$_2$ attributions) and attributions of survival value (function$_3$ attributions). For that reason I start my discussion of the meaning of survival value (section 6.2 to 6.5) with an elaborate account of the differences between these two kinds of function attributions (section 6.2). In section 6.3 I discuss the main objections raised against the use of counterfactual comparisons to single out functions. I argue that although some of these criticisms apply to the use of counterfactuals to determine causal roles (functions$_2$), they do not apply to my use of counterfactual comparisons to determine survival value (function$_3$). In section 6.4 I discuss the notion of adaptation and its relation to the different notions of functions I distinguish.

In the second part of this chapter (section 6.5-6.6) I discuss the attempts of Canfield (1964) and Horan (1989) to account for the explanatory use of appeals to survival value. According to Canfield appeals to survival value merely show what an item is good for. I argue by means of examples that this account neglects much of what is achieved by appeals to survival value. According to Horan appeals to survival value explain how a certain trait is maintained in the population. She contends that such explanations employ so-called “consequence laws”. These laws state that a certain trait arises because it has beneficial consequences. I argue that the notion of consequence laws is confused but that the idea that appeals to survival value are used to explain the maintenance of a certain trait is promising. Sober (1984) works out this idea five years before Horan published her paper. Reeve & Sherman (1993) present a similar idea. I agree with these authors that in many cases appeals to survival value inform us why a certain trait is maintained in the population. However, appeals to survival value also provide insights in the way in which an individual hangs together. This latter achievement is ignored by the accounts of Sober and Reeve & Sherman.

In section 6.7 I draw the conclusion that the current proposals within the survival value approach fail to account for the insights gained by appeals to survival value in design explanations.

Part I: The meaning of ‘survival value’

6.2 Attributions of causal roles and claims about survival value

6.2.1 The definition of ‘function’ according to the survival value approach

Proponents of the survival value approach tend to think of function attributions as having the following standard form:

the / a function of X is to do Y

where:

X can be anything of a living organism: an item, a behaviour, a trait, the structure of an item, the activity of an item and so on, and

Y is an activity of X or an effect of X.

The survival approach explicates function as a special kind of activity or a special kind of effect, namely as those activities / effects that enhance the fitness of the organisms that perform that activity or produce that effect. This means that the analysis of function within the survival value approach is something like this:

> the function of X is to do Y if and only if Y enhances the fitness of the individuals that have / perform X.

The notion of 'fitness enhancing' is comparative and typically explicated by means of a counterfactual: an effect enhances the fitness of an organism if the fitness of that organism would be lower it that effect did not occur.

For example, Canfield (1964) defines 'function' as follows:

> a function of I (in S) is to do C *means* I does C; and if, *ceteris paribus*, C were not done in an S, then the probability of that S surviving or having descendants would be smaller than the probability of an S in which C is done surviving of having descendants (Canfield 1964: 292)

where:

I is an item,
C an activity, and
S a species.

Wimsatt (1972) argues that function attributions have the following form:

> According to theory T, a function of behaviour B of item i in system S in environment E relative to (purpose, standard or perspective) P is to do C (Wimsatt 1972: 42).[1]

In the case of natural organisms the relevant theory is evolutionary theory and the purpose or standard is something like survival, reproduction, increase of fitness or selection. Roughly spoken, a function of an activity of an item (of a living organism) is an effect that increases an organism's life chances:

> an entity could be regarded as functional if its presence or operation produced an increase in the probability of purpose-attainment (Wimsatt 1972: 55).

Wimsatt notes that this criterion is comparative:

> an increase in the probability of purpose-attainment must be an increase relative to some other state of affairs (Wimsatt 1972: 55).

The determination of a unique reference state is, according to Wimsatt, an important unsolved problem in the philosophy of biology.

---

[1] Note that on Wimsatt's account it is the "behaviour" (i.e. activity) of an item that has a function not the item itself.

Twelve years later, Bigelow and Pargetter (1987) maintain that

> something has a (biological) function just when if confers a survival-enhancing propensity on a creature that possesses it (Bigelow & Pargetter 1987: 192)

They say that there are several ways to spell out the notion of a “survival-enhancing propensity”:

> [Propensity] 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 (Bigelow & Pargetter 1987: 194).

The phrase “increasing the probability” suggests that function statements are comparative and the phrase “what it would have been” suggests that Bigelow & Pargetter want to determine function on the basis of counterfactual comparison.[2]

> A character or structure[3] has a certain function when it has a propensity for selection in virtue of that character or structure having the relevant effects (Bigelow & Pargetter 1987: 194)

Horan (1989), too, talks of ‘fitness enhancing’ in connection with function:

> One can say that questions about the function of a given pattern of social behaviour are a way of asking how that behaviour enhances the fitness of an individual who engages in it (Horan 1989: 135)

She does not spell out this notion.

None of the proponents of a survival value approach distinguishes explicitly between attributions of causal roles (function$_2$) and claims about survival value (function$_3$). In the remainder of section 6.2 I argue why it is important to make such a distinction. I also show that the kind of wording the proponents use to define function indicates that they confuse causal roles and survival value.

6.2.2 A survey of the differences

In section 2.2.3 I listed the main differences between attributions of causal role and claims about survival value. Let me briefly review them.

---

[2] Later, on the same page they draw the conclusion that “a character or structure has a certain function when it has a propensity for selection in virtue of that character or structure having the relevant effects” (Bigelow & Pargetter 1987: 194). As natural selection is always among real variants this quote suggests that Bigelow & Pargetter think of functions as effects that confer an advantage in the competition with real variants. However, on that construal traits that do not vary will have no function at all.

[3] Bigelow & Pargetter’s “character” is roughly equivalent to my “trait” and “structure” to my item.

First, causal roles are typically attributed to an item or behavioural pattern whereas claims about survival value concern traits such as the presence or character of that certain item or behavioural pattern. For example, Schwenk (1994) claims that the snake's tongue (an item) has a causal role in trail-following and that the forked character of this item has survival value (example 2.3 of section 2.2.2). Similarly, it is said of a certain behaviour that it has a causal role in say feeding or territory defence and of the character of that behaviour (e.g. flock feeding, or threat display) that it has survival value (example 3.2 and 3.3 of section 2.2.3).

Secondly, attributions of causal roles are concerned with how an item or behaviour fits into an organism's machinery, claims about survival value are concerned with more global effects such as an item's contribution to the survival, reproduction or fitness of an organism. The causal role of the heart to pump the blood around can be established independent of the survival value of that latter activity. It suffices to show that the heart contributes to circulation by pumping. On the other hand, when studying the survival value of the egg shell removal behaviour Tinbergen and his colleagues (1962) were interested in the effect of that behaviour on the reproductive success of the organisms that perform it (example 3.1 of section 2.2.3). Similarly, to establish that the social foraging behaviour of black headed gulls has survival value because it allows each individual gull to catch more fish than it would have done otherwise, one must assume that the number of catches correlates with fitness (example 3.2 of section 2.2.3). As a result of this difference, it is appropriate to talk of the causal role of an item or behavioural pattern *in* an organism (or *in* a system of that organism) and of the survival value of the presence or character of that item or behavioural pattern *for* the organisms that have it.

Note, that the above distinction between the notions of function as causal role and function as survival value explains how it is possible that there are cases in which one speaks of an item or behaviour as having a function even though its activity is neutral or detrimental to survival. Robert Cummins (1975: 755/6) observes that we would say that the function of the wings of pigeons is to enable them to fly as long as they do so, even if flying ceased to contribute to survival. This argument may seem far fetched and in fact Cummins does not substantiate it. Yet, among biologists it is well known that flight is deleterious to birds living on oceanic islands because they are liable to be blown to the sea (Lack 1947) but this does not preclude them from analyzing the muscles of such birds in terms of their contribution to flight. In section 4.2.2 I mentioned another example: the glandular hairs on the leafs of sundew are said to have the function to catch flies, even in circumstances in which catching flies does not contribute to survival. In such cases 'function' clearly means causal role (and only causal role).

Thirdly, attributions of survival value are relative to a certain environment, attributions of causal roles not. Although it depends on the environment whether an item is capable of performing its causal role, it does not depend on the environment whether it *has* that causal role.[4] Consider the examples given in section 2.2.2 (the causal role of the heart, thymus and the snake's tongue). In each case the causal role of the item (to pump blood, to initiate the differentiation of T-lymphocytes and to sample chemicals when following scent trails) is singled out by a capacity explanation and an item has that causal role independently of the environment in which the organism actually lives. The causal role of the heart in the circulatory system is to pump blood if that is the way in which the heart contributes to the circulatory system in environments in which that systems works. On the other hand, the survival value of the egg shell removal behaviour in birds (example 3.1 of section 2.2.3) depends on the presence of predators and their visual capacities, that is on the environment of the organisms. In there were no predators the behaviour would not have survival value. The survival value of social foraging, too, depends on the environment: if the gulls were fed with dead fish, social foraging would not have survival value.

Fourthly, perhaps the most important difference is the following: attributions of survival value are essentially comparative and attributions of causal role not. When biologists say that the heart pumps the blood around they say something about the causal role of that item in organisms with hearts. However, when they say that social foraging has survival value because it allows each gull to catch more fish than it would have done otherwise, they compare one type of behaviour (fishing in flocks) with another type of behaviour (fishing alone). As I discussed in section 2.2.3, the comparison is often counterfactual: the real organism is compared to a hypothetical organism that is similar to the real organism but lacks the relevant trait (or possess that trait to a lesser degree).

6.2.3 0Example: respiration

I will now illustrate the differences between causal role and survival value by means of a number of examples. My first example is about respiration.

In adult tetrapode vertebrates (amphibians, reptiles, birds and mammals) respiration is typically performed by means of lungs[5], whereas in fishes respiration is typically performed by means of gills.[6] The respiratory role of those organs can be established independent of the survival value of respiration. It suffices to show that they contribute to respiration. The respiratory system in its turn is part of the system concerned with the exchange of respiratory gasses (oxygen and carbon dioxide) between the external environment of the organism and its metabo-

---

[4] If an item does not perform its role it is possible to determine what its role is by studying organisms that are build in a similar way.

[5] In many amphibians the skin too has an important role in respiration.

[6] Some fish use lungs.

lizing cells. These complex tasks can be divided into three sub tasks: exchange of respiratory gasses between the external environment and the circulatory system, transport of the gases, exchange between the circulatory system and the cells. These three tasks are performed by two interconnected systems: the respiratory organs (in which oxygen is taken from the environment and carbon dioxide released in the environment), and the circulatory system (which transports oxygen from the respiratory organs to the cells and carbon dioxide back from the cells to the respiratory organs, exchange of gases between cells and circulatory systems takes place in the capillaries, which are parts of the circulatory system).

Given the causal role of the lungs and the gills, two kinds of questions about survival value arise. The first is concerned with the survival value of the respiratory system as a whole. The second with the specific character of the organs that perform the respiratory task. The need for fulfilling a respiratory task is implicitly discussed in chapter 4 (section 4.2.3). There I explained that due to the limits of diffusion the volume of an organism demands for circulation. It is obvious that the demand for performing a respiratory task arises together with the development of a circulatory system. A further question is why this task is concentrated in specialized organs. The answer is that there are many disadvantages to the use of the entire body surface for respiratory exchange. It follows from Fick's law of diffusion (discussed in section 4.2.3) that to maintain a diffusion rate large enough to fulfil the organism's needs (i) the distance across which the gas must diffuse must be as small as possible, (ii) the surface available for diffusion must be large enough, and (iii) the material across which the gas diffuses must be readily permeable to that gas. A thin skin, however, is easily damaged and a skin that is easily permeable to oxygen and carbon dioxide is also easily permeable to water (which is a severe disadvantage on land and in aquatic environments with an osmotic pressure that differs from the organism), enlargements of the outer surface of the body would disturb the stream line and so on.

An example of a question about the specific character of the respiratory organs is the question why respiration in water is performed by means of gills and respiration in air by means of lungs. To answer to this question, biologists compare the physical qualities of water with those of air. The concentration of oxygen in air-saturated water is about 1/30 of that in air and the rate of oxygen diffusion in air is about 30 000x higher than in water. As a result, aquatic animals need to ventilate a vastly larger volume than land animals to extract the same amount of oxygen. Air breathing faces other difficulties: there is the continuous risk of desiccation and there are the problems caused by gravitation. As Archimedes testifies these problems do not occur in water. The main differences between gills and lungs are explained by pointing to these physical differences. The flow of water across the gills is unidirectional. This increases the efficiency of ventilation. Compared to lungs, gills have a much larger respiratory surface with a much thinner membrane. This compensates for the smaller difference of the concentration in and outside the membrane. Such a structure would not fit for air breathing. Due to the problems of gravitation the immense increase of the surface area in gills would not be possible on land: the finely divided and thin filaments collapse against each other. Lungs are internal, which reduces the risk of desiccation and provides the means for structural support to counteract gravitational effects. In lungs the flow of air is bi-directional: there is an inhalation and an expiration phase.[7] This is much less efficient than in gills (recall that air breathing imposes lesser demands on ventilation) but it reduces the loss of water. The pumping mechanism in tetrapodes has much less power than in fishes. For these two reasons, lungs would not work in an aquatic environment.

Let us compare the differences between asking for the causal role of lungs and gills, and the questions about their survival value. First, the causal role of respiration is attributed to certain organs (lungs and gills), whereas the questions about survival value concern the presence of these organs and the way in which they are built.

Second, one can say what the causal roles of lungs and gills are (pointing to these organs) without saying anything about their survival value. Indeed, the causal role of lungs and gills as respiratory organs was established long before the survival value of having a respiratory system. The causal role became clear due to the work of Priestley and Lavoisier in the second half of the eighteenth century. The survival value of having a respiratory system became clear due to the work of Krogh in the 1940s.

Third, to say that lungs and gills have a causal role as respiratory organs is to say something about the way in which those organs fit into the organism’s structure of capacities. It tells us which tasks those organs perform, but not how performing these tasks influences the life chances of the organisms having those organs. The accounts of their survival value are concerned with the (dis)advantages of certain ways of performing that causal role for the organisms as a whole. This is clear in the account of the survival value of having a respiratory organ: this account points out that vertebrates that lack a respiratory organ would have certain disadvantages. The explanation of why aquatic gas exchange is done by means of gills and air breathing by means of lungs, too, points to the advantages and disadvantages for the organism as a whole: it points out that animals that extract oxygen from the air by means of gills and animals that extract oxygen from water by means of lungs, would not get enough oxygen to survive.

Fourth: to say that lungs have a causal role in respiration or in air breathing is to say something about the way those organs are used in certain organisms, full stop. However, to say that lungs are better for air breathing than gills, is to say something about how lungs fare in comparison to gills. The survival value of having a specialized organ for respiration too is of a comparative nature: it compares organisms having a specialized organ with hypothetical organisms using the whole outside of the body for respiration. Finally, whether lungs and gills have a

---

[7] In birds the flow is virtually unidirectional. This meets the increased demand for oxygen imposed on them by their flying lifestyle.

causal role as respiratory organs is independent of the environment. Of course, it depends on the environment whether lungs and gills are able to perform that causal role, but whether they have that causal role depends on the way in which that organism is organized. On the other hand the extent to which those organs have survival value depends on the environment. Gills have more survival value in water, whereas lungs have more survival value in water. A specialized organ has survival value in environments in which the oxygen concentration is low.

6.2.4 Example: the heart

Most proponents of a survival value approach mention the statement ‘the heart propagates the blood’ (or something like that) as an example of an attribution of function that fits their analysis of function statements as statements that tell us how a trait contributes to survival. In this section I argue that this attribution is on the contrary a typical example of an attribution of a causal role.

We can attribute the blood-propagating function to the heart without any knowledge of the survival value of having a circulatory system. Indeed, that was what happened during the history. Harvey (1628) voiced the causal role of the heart as a pump long before Krogh (1941) demonstrated the survival value of circulating blood. The attribution is concerned with the local effect of the heart on a containing system, not with the global effect of the heart on the life chances of organisms with a heart. The attribution is not comparative and the function of the heart as a pump is not relative to the environment.

To clarify this point let us see what questions about the survival value of the heart would look like. There are two kinds of questions about the survival value of the heart: one concerns the survival value of having a circulatory system, the other concerns the specific character of the heart (given its causal role). The survival value of having a circulatory system was elaborately discussed in chapter 4 (section 4.2.3). The need for a circulatory system was established by comparing the actual organisms with a hypothetical spherical organism larger than 1 mm^3 that had to rely on diffusion alone. This consideration establishes the survival value of a system of convection (in addition to diffusion): it enables organisms to grow larger.One may want to compare different systems of convection in a further study. Insects do not transport oxygen in the blood and they do not have a localized respiratory organ. Instead, they use a system of tubes (called ‘trachea’) invaginated from the outside. This system seriously limits the size attainable by insects. Given a system of blood circulation one may compare the survival value of different kinds of pumps. Compare, for instance, a muscular pump with a ciliate pump or compare a localized pump with a system of beating blood vessels.

Someone might argue that the attribution to the heart of the function to pump the blood is comparative: it compares an organism with a heart with an organism without a heart. The attribution tells us that organisms having a heart survive better than organisms without a heart because the heart pumps. This argument confuses the two kinds of questions about survival value. Questions about the survival value of an item depend on a preceding attribution of a causal role to that item in maintaining a certain capacity. If that causal role is known, one may ask about the survival value of having that capacity (in comparison to lacking that capacity) (type 4a questions) and about the survival value of the way in which that causal role is performed (in comparison to other possible ways of performing that causal role) (type 4b questions). It is not clear what is meant by the survival value of having a heart as compared to lacking a heart (more specifically it is not clear what is meant by ‘having a heart’). Is the actual organism compared with an organism in which the circulatory task is performed by means of another pump or is it compared with an organism in which the circulatory task is not performed? In the first case the survival value of having a heart is not that it pumps but, for example, that it pumps more efficiently than that other pump. In the second case the survival value of the heart is that it makes it possible that the body is larger. In neither case has the heart survival value because it pumps.

6.2.5 Example: territory behaviour

The two examples above were concerned with morphology. To show that the distinction between causal role and survival value is not peculiar to that discipline, I now discuss an example from ethology. My example concerns territory behaviour. When ethologists say that a certain type of behaviour (characterized by its form) in certain organisms has as its function the maintenance of a territory they say something about the causal role of that behaviour. For example, when Catchpole (1979: 31-37) says that singing in birds has a function in territory behaviour, he means that singing contributes to the capacity to maintain a territory (see section 2.3.2). Similarly when Habibi and his colleagues (1993) say that the habit of sand gazelles to urinate at certain places and to leave secretes from the preorbital glands at certain places both have a function in maintaining a territory they mean that leaving such scents contributes to maintenance of a territory (example 3.2 of section 2.2.3). Such statements do not inform us about the survival value of territory behaviour. They inform us neither about the survival value of maintaining a territory (as compared to not doing so) nor about the survival value of one type of behaviour (for example singing) as compared to other means to maintain a territory) for example conspicuous colouring). The behaviours I mentioned have a causal role in territory behaviour even if the maintenance of a territory does not contribute to survival. The attribution is concerned with the effect on a local context (its effect on maintaining a territory) and not comparative.

Given this causal role one may ask questions about the survival value of certain types of territory behaviour as compared to other types of behaviour. This is, for example, what Habibi c.s. do. They compare the survival value of leaving scent marks with that of leaving dung piles and suggest that scent marks work better when the gazelles have relatively small territories occupied during a relatively short period and dung piles work better when the territories are bigger and maintained during the whole year.

The maintenance of a territory itself may have different causal roles. One is to provide a reliable food supply, another to find mates. Two kinds of questions about survival value arise: ‘under what circumstances is holding a territory a suitable manner to fulfil these causal roles and why is that so?’ and ‘why is the manner in which the territory is maintained a good way to perform that causal role given the environment and the way the organism is built and behaves’. The second question is answered by Habibi c.s. by comparing the manner in which sand gazelles and mountain gazelles maintain their territory. As I said, they argue that if reproduction takes place during a short season (as in sand gazelles) it is better to round up females in harems and protect them vigorously against rivals, whereas if reproduction takes place during the whole year (as in mountain gazelles) it is more useful to have large territories and defend them with help of threat displays.

6.2.6 Example: egg shell removal in birds

My final example concerns the study of Tinbergen and his students (1962) of the egg shell removal behaviour in black-headed gulls (example 3.1 of section 2.2.3). As in the case of Harvey and Miller, Tinbergen c.s. accomplish several things at once. Their main concern (at least in the parts of the paper that deal with the survival value of the egg shell removal behaviour) is the question why it is useful to remove the empty egg shell rather than to leave them were they are. To answer this question they put forward the hypothesis that this behaviour has a causal role in the anti-predator system, namely that it helps to maintain the camouflage of the brood. Next, they show that this role is better performed if the empty shells are carried away than if they are left near the nest, because carrion crows and herring gulls find the nest more easily if there lays an empty egg shell in its immediate environment. Tinbergen c.s. conclude that the results of their experiments “leave little room for doubt about the survival value of egg shell removal as an anti-predator device”. This is only a poor summary of what they have shown, namely (1) what the causal role of the egg shell removal behaviour is (egg shell removal has a causal role in the anti-predatory system, namely to camouflage the eggs), and (2) why that behaviour is useful (it has survival value to remove the empty egg shell after the chick has hatched rather than to leave it where it is because that helps to maintain the camouflage of the eggs).

The first claim (egg shell removal has a causal role in the anti-predatory system, namely to camouflage the eggs) is a claim about the way in which the behaviour contributes to the particular system of the animal that has the task to avoid being consumed by other animals. The second claim is concerned with the reproductive success of the animals that perform this behaviour. Note that a contribution to the anti-predator system does not, by definition, contribute to survival. A small animal that spends the day hiding for predators might starve because it cannot reach its food. An animal that runs away from one predator, might run into the arms of another. A behaviour that is useful to avoid one kind of predators might give other kinds a chance (recall that some bleak-headed gulls specialize in eating the eggs and chickens whose parents have left the nest). This shows that it makes sense to distinguish between statements that position a certain behaviour in a system that performs a certain task (attributions of a causal role) and statements that explain why it is useful to perform the task in the way it is performed (claims about survival value).

As in the other examples, the attribution of a causal role in this example is not relative to a certain environment. Of course, it depends on the environment whether the egg shell removal succeeds in maintaining the camouflage of the egg but it does not depend on the environment whether egg shell removal has that causal role. On the other hand it does depend on the environment whether egg shell removal has survival value: if the gulls live in an environment in which there are no predators, such as carrion crows, that can find a nest with an empty egg shell near to it more easily than a nest without such an empty shell, then the removal behaviour has no survival value (but it still has the causal role to maintain the camouflage). If the birds live in an environment in which there are no carrion crows and egg shell removal makes it easier for other predators to find the remaining eggs the removal behaviour may be even detrimental to reproductive success.

Finally, the attribution of the causal role to maintain the camouflage is not comparative. The behaviour has this causal role no matter what effects other kinds of behaviour would have. The truth of the claim that the removal behaviour has survival value on the other hand depends on the behaviour which with it is compared. Tinbergen c.s. show that the egg shell removal behaviour has the causal role to maintain the camouflage by means of an experiment in which they compare nests with normal eggs and nests with painted eggs; and also nests with an empty egg shell and nests without an empty egg shell. This might suggest that the relevant attribution of a causal role is comparative. However, one should not confuse the content of a certain statement with the evidence for that statement. If one shows that C does not occur when one replaces *A* by *B*, one has provided evidence for the assertion that *A* contributes to *C*. However, ‘*A* contributes to *C*’ does not mean the same as ‘A effects C better than *B*’. Similarly, one may show that a certain item has a causal role in performing a certain task by replacing that item with another one and studying the effect of this replacement on the performance of the task. However, this does not imply that the attribution is comparative.

6.2.7 Conclusion

In section 2.2.2 I stated that attributions of causal roles have the following basic form:

> item / behaviour *i* has causal role *f* in maintaining activity / capacity *c* of system *s*

for example: the heart of vertebrates (*i*) has the causal role to pump blood (*f*) in maintaining the capacity of the circulatory system (*s*) to circulate blood (*c*). This might be reformulated as:

> the function$_2$ (causal role) of item / behaviour *i* in maintaining activity / capacity *c* of system *s* is to do *f*

for example: the function$_2$ (causal role) of the heart of vertebrates in maintaining the circulatory system's capacity to circulate blood is to pump blood.

In section 2.2.3 I said that claims about survival value have the form:

> trait *t* has survival value in comparison with trait *t*' under conditions *c* because ... (follows an explication of why an organism in condition *c* is better off if it has *t* rather than *t*' — the explication is often that a certain causal role *f* is performed better by an organism with *t* than by an organism with *t*').

for example: using lungs (*t*) rather than gills (*t*') to respire (*f*) has survival value for animals that breath air (c) because it solves the problems of desiccation and gravitation that would occur if they used gills. This claim is not easily converted into a statement of the form 'the function of ... is ...'.

In section 6.2.3–6.2.6 I have substantiated these two claims by means of examples.

In section 6.2.1 I stated that the proponents of the survival value approach tend to explicate the meaning of 'function' as follows:

> the function of item / behaviour / character / trait *X* is to do *Y* if and only if *Y* is an activity / effect of *X* that enhances the fitness of the individuals that have / perform *X*.

The very wording of this definition betrays a confusion of attributions of causal role and claims about survival value. The definition says that function statements might concern almost anything: items, behaviours, characters, traits, and so on. This assumption ignores the first difference between attributions of causal roles and claims about survival value: causal roles are attributed to items or behaviours, whereas claims about survival value are concerned with traits (i.e. the presence or character of an item or behaviour). What is more important, the analysandum is non-comparative (as are attributions of causal role), but the analysandum suggest a comparison (as with claims about survival value). The non-comparative character of the analysandum suggests that the analysis aims to define function as causal role. The same suggestion is conveyed by the 'the function of ... is ...' form of the analysandum (as I said above, claims about survival value are not easily converted to this form). However, the phrase 'enhances the fitness' suggests that the function statement compares the effects of the item or behaviour in study with possible alternatives that might replace that item or behaviour (as with function as survival value). In other words, the survival value approach appears to analyze the notion of *function as causal role* in terms of fitness enhancing effects, and, hence, in terms of counterfactual comparison. This is a severe confusion (as I have shown in section 6.2.2–6.2.3). It is the notion of function as survival value that should be explicated in terms of counterfactual comparision. The notion of function as causal role is better analyzed along the lines of Cummins .

6.3 Functional counterfactuals

6.3.1 Introduction

There are many ways to test hypotheses about survival value. Roughly speaking, the evidence is of three kinds: calculations, experiments and correlations. All these methods aim to establish that under certain external and internal conditions the fitness of the organisms that have a certain trait is greater than the fitness of organisms that lack that trait or have another trait instead. The methods to provide evidence for survival value are worth a special study and I'm not going to delve into that question deeply. However, there is one issue that deserves special attention. This is the issue of counterfactual comparison. As I said attributions of survival value are always comparative. In the case of correlational studies one compares real organisms that live in different conditions. However, if the evidence is provided by calculations or experiments a comparison is made between a real organism and a hypothetical one. The hypothetical organism is similar to the real organism but the trait of which the survival value is studied is absent or present to a lesser degree. The results of such a comparison are expressed by means of a statement I call a 'functional counterfactual' (section 2.2.3). Functional counterfactuals state that if an organism lacked the trait under study (or had it to a lesser degree) it would survive or reproduce less well.

Many philosophers find such counterfactual judgements highly suspicious. For example, Frankfurt and Poole (1966) denounce the way in which Canfield (1964, 1965) employs counterfactuals in his account of function as "biologically meaningless" (p. 72). More recently, Millikan writes about the reference to counterfactuals in the account of function presented by Bigelow and Pargetter (1987):

> Unfortunately, exactly in this sort of context, counterfactuals are notoriously indeterminate in truth value (Millikan 1993a: 39).

Roughly speaking, three kinds of objections have been made against an analysis of survival value in terms of counterfactuals: (1) the counterfactual condition runs against the laws of nature, (2) the counterfactual condition is not well defined, (3) counterfactual judgements are too complex to be made in a reliable way. The issue is complicated by a tendency to confuse attributions of causal role and statements about survival value. As I will show, both advocates and opponents of a counterfactual analysis of ‘function’ are concerned with the use of counterfactuals to distinguish causal roles from mere effects. I defend the use of counterfactuals to assess survival value. That is, I show that the arguments brought up by the opponents against the use of counterfactuals to distinguish causal roles from mere effects, do not count against the use of counterfactuals to assess survival value. I start with a summary of Canfield’s (1964, 1965) account of function. This is not only the first account that explicitly employs counterfactuals but also the most elaborate discussion of that use. Next, I show that functional counterfactuals are central to explanations that appeal to survival value (i.e. design explanations). Then I deal with the three objections mentioned above.

6.3.2 Canfield’s account of function attributions

One of the first statements of a survival value account of function attributions is Canfield’s “Teleological explanation in biology” (1964). Canfield defines function attributions (he calls them “functional analyses”) as sentences that contain “expressions like ‘function’, ‘purpose’, ‘role’ and ‘in order to’” (p. 285). They can be rewritten in the standard form “a function of ... is to ...”. Examples are: “a function of the liver is to secrete bile”, “the heart beats in order to circulate blood”[8], and “a function of the thymus is to produce lymphocyte cells”. Function attributions serve as explanations if they are offered as an answer to a request for explanation. For example, a function attribution as “the heart beats in order to circulate the blood” serves as an explanation if it is offered in answer to the question “why does the heart beat?”.

Canfield starts his analysis of the meaning of function attributions with a real example: the study of the function of the thymus as described by Burnet (1962) in the *Scientific American.* I have summarized Burnet’s paper in section 2.2.2 (example 2.2). For many years biologists searched in vain for an effect of the removal of the thymus on the physiology of adult organisms. Study of new-born mice revealed that the thymus has a function in the development of the immune system: it initiates the differentiation of T-lymphocytes. According to Canfield this example shows that function attributions describe what an item does that is useful for the organism to have:

---

[8] Unfortunately Canfield does not state explicitly how to transform the statement ‘the heart beats in order to circulate blood’ into his standard form. This transformation is problematic because the original statement relates three things (namely an item (the heart), an activity of that item (beating) and the capacity in which that item has a role (circulating the blood)) and the standard form only two (an item and a role).

> the example indicates that we give a functional analysis of the thymus in, e.g. the mouse, if we state what the thymus does that is useful to the mouse. If, e.g. we have shown that in the mouse the thymus produces lymphocyte cells which themselves have such and such useful effects, then we have shown that a function of the thymus is to produce these cells (Canfield 1964: 287).

In his next section Canfield states (without appeal to this example and indeed without any further argument) that the activity of an item can be useful in two ways: it may be useful by increasing the life chances of the organisms having that item or by increasing their probability of having progeny (p. 291). In both cases a normal organism is compared to a non-normal organism in which the function is not performed but which is the same as the normal organisms in all other relevant respects. This can be stated more formally by means of a subjunctive conditional. Let *I* be an item occurring in (all (?) organisms of?) species *S* and let *C* be an activity of that item, then

> a function of *I* (in *S*) is to do *C* *means I* does *C*; and if, *ceteris paribus*, *C* were not done in an *S*, then the probability of that *S* surviving or having descendants would be smaller than the probability of an *S* in which *C* is done surviving of having descendants (Canfield 1964: 292).

The *ceteris paribus* clause assumes that (i) the two organisms are alike in all relevant aspects other than doing *C* and otherwise normal, and (ii) the two organisms are compared with respect to the same (normal) environment.

Note that Canfield confuses attributions of causal roles and attributions of survival value. His examples concern attributions of causal roles. However, his explication of what it is to be useful is in terms of survival value. This is the reason why he fails to support his analyses of what it is to be useful with his examples.

Lehman (1965b) puts forward the following counterexample to this analysis:

> A function of the heart (in human beings) is to produce a pulse (Lehman 1965b: 327).

In response, Canfield (1965) clarifies the use of this subjunctive conditional with help of the notion of an “action-sentence chain”. An action-sentence is a sentence that describes what a certain item does, examples are ‘the liver secretes bile salts’ and ‘biotin participates in pyruvate metabolism’. Two or more action-sentences may be combined into an action-sentence chain on the following terms: (1) the first sentence in the chain says that some *A* does *B*, (2) the second sentence says that *B* does something else, *C*, (3) the third sentences (if any) says that *C* does *D*, and so on. An example of an action-sentence chain is: <‘the liver secretes bile salts’, `bile salts emulsify fats in the intestine’>. Now, the requirement that the functional organism *S1* be compared with another organism *S2* that is “alike in all respects other than *C*” could be worked out as follows. Suppose one has a description of *S1* in terms of its action-sentence chains. Then *S2* is another organism of the same species whose description contains all action-sentence chains of *S1*'s description with the exception of those chains in which *C* is mentioned:

> To suppose that *S1* is like *S2* except that in *S2 C* is not done is to suppose that *S1* is characterised by all true sentence chains, whereas *S2* is characterised only by a subset of the true sentence chains. That is, *S2* is not characterised by any sentence chain in which 'C' occurs. The description of *S2* which results is that of an animal in which we have somehow removed the effect *C*, and also every result or effect of *C*, results of those results, and so on (Canfield 1965: 330).

On this account pumping is a function of the human heart since removing all effects of pumping from a human organism results in an organism that is not viable due to a lack of capacity to transport nutrients to the cell. On the other hand, producing a pulse is a mere effect and not a function of the human heart since removing all effects of pulsation does not affect the chances of survival (because transportation is not achieved by means of the pulse *S2* would transport nutrients just as well as *S1*). Canfield adds to this:

> My point might be put more intuitively as follows. Suppose we could add to an animal's heart an instrument which resulted only in making the animal's blood flow at a steady state, rather than in spurts. So far as we know this change would neither raise nor lower the animal's chances of survival.[9] Since this is true, Lehman's sentence *P2* ["the heart produces a pulse in human beings and if *ceteris paribus* a pulse were no produced in a human being, then the probability of that human being surviving or having descendants would be smaller that the probability of a human being in which a pulse is produced surviving of having descendants"] is false, rather than true and thus *P1* ["a function of the heart in human beings is to produce a pulse"] is no counterexample to my analysis (Canfield 1965: 330).

---

[9] Actually, Canfield is wrong about this. As a simple calculation shows, dampening the pulse has survival value and that's why there are pulse dampeners in almost any circulatory system ever evolved (in vertebrate blood circulation the dampening is performed by the elastic walls of the central arteries). Compare a hypothetical system in which the blood flows in spurts with a hypothetical system in which the blood flows continuously. Both systems circulate the same amount of liquid in the same time. Putting friction aside, the energy needed to transport this mass is equal to the kinetic energy of the liquid. If a mass *m* is transported at a steady pace *v* the energy needed to transport this mass equals $^1/_2 mv^2$. In the system with intermittent flow the time available for circulating the same amount of liquid is half that of the continuous system (half of the time there is no flow). Therefore, the velocity is twice and the energy used is four times that of the continuous system. This shows that if the blood flowed in spurts the energy needed to transport a certain mass of fluid would be four times that needed to transport the same amount in a continuous manner.

6.3.3 Functional counterfactuals in biology

I gave several examples of the use of counterfactual statements in explanations in section 2.2.3 and 4.2.3. These examples show that Canfield is right that counterfactual comparisons do play a basic role in explanations in functional biology. This is, for example, how Tinbergen and his colleagues explain the colour of the eggs of the black-headed gull:

> The natural egg colour of the Black-headed Gull’s eggs makes them less vulnerable to attack by predators hunting by sight than they would be if they were white (Tinbergen et al. 1962: 80/81).

The following counterfactual explains the removal of empty egg shells:

> Egg shells would endanger the brood if they were not carried away (Tinbergen et al. 1962: 82).

Other counterfactual statements used in explanations discussed in previous sections:

> Scent-marks may be less long-lasting, and it would not be possible for a male to replenish them fast enough to be effective in a large territory (Habibi, Thouless & Lindsay 1993: 51).
>
> An extended birth season is disadvantageous to a migrating herd as the neonates and lactating females would be under stress when travelling long distances to new feeding grounds (Habibi, Thouless & Lindsay 1993: 52).
>
> Oxygen would probably not diffuse into them [flatworms] fast enough for their requirements if they were much thicker than this (0.5 mm) (McNeill Alexander 1975: 158).
>
> If the blood [of *Helix*] did not contain haemocyanin the tissues could not be supplied with oxygen at the required rate unless the heart were larger or beat faster (McNeill Alexander 1975: 276).
>
> An earthworm more than about 30 mm in diameter would not be feasible unless it had a lower metabolic rate [..] or the blood came nearer the surface of the body (McNeill Alexander 1975: 356)

A quick look in any textbook of functional biology would yield many more examples. I add just three. William Keeton and James Gould use a functional counterfactual to explain why plants in contrast to animals do not need special gas-transporting organs:

> Most of the intercellular space in the tissues of land plants are filled with air, in contrast to those in animal tissues, which are filled with fluid. [...] Since oxygen can diffuse some 10,000 times faster through air than through liquids, the intercellular air-space system ensures that all cells [...] are adequately supplied. If the oxygen had to diffuse through liquid from the surface of a plant organ, it would penetrate less than one millimeter, and all of the more internal cells would be deprived of oxygen and could not respire (Keeton & Gould 1993: 800).

This is their explanation of why fish actively move water across their gills:

> If the water remained still, the O_2 in the vicinity of the exchange surfaces would soon be depleted and it would not be renewed by diffusion fast enough to sustain the animal (Keeton & Gould 1993: 804).

At the same page they explain why the blood stream and the water stream in the gills are in opposite directions:

> In short, this countercurrent strategy [...] maximizes the amount of O_2 the blood can pick up from the water. This would not be the case if the two fluids had the same direction of flow (Keeton & Gould 1993: 804).

These examples show the importance of counterfactual comparison in explanations in functional biology.

6.3.4 Functional counterfactuals and laws of nature

Frankfurt and Poole object against Canfield's (1964, 1965) analysis that functional counterfactuals run against the laws of nature:

> If a certain activity is performed in one specimen and not in another, it is impossible that the two specimen should differ only in this. An organism's activities arise out of its bodily structure, and its structure must be altered in order to change its activities. The notion of two organisms which have the same structure, but in which different activities take place, violates our ideas of causality (Frankfurt & Poole 1966: 71).

They add to this:

> Suppose we wish to know whether secreting bile is a function of the liver in mice. Canfield advises us to examine a mouse whose liver does not secrete bile but which is otherwise identical with a mouse whose liver does secrete it. There can be no such mouse. If it were built the same as other mice, it would naturally secrete bile just as other mice do (Frankfurt & Poole 1966: 71).

In a similar vein, Nissen (1970) argues that Canfield's (1965) clarification does not dispel Lehman's (1965b) counterexample. The main problem is that

> In order to remove one of the effects without bringing in other mechanisms, the cause must be removed (Nissen 1970: 194).

This poses a dilemma to Canfield's account. In order to determine the survival value of the pulse one should either remove the cause (the heart) or add an additional mechanism (such as a pulse dampener). Both horns yield counterexamples. Removing the heart has a negative effect on survival. This would incorrectly allocate a function to the pulse. Adding additional mechanisms violates Canfield's *ceteris paribus* clause. Modification of this clause so as to allow for additional mechanisms such as pulse dampeners would open the door to a host of other counterexamples. For example, if one allows pulse dampeners to determine whether producing a pulse is a function of the heart, one must also allow a plastic skin covering and modification of the lungs to determine whether cooling the skin is a function of perspiration. However, if the lungs are modified in such way that panting cools the body as efficiently as the skin than removing perspiration with help of a plastic skin covering will have no effect on survival. Hence, on this construal cooling the skin is not a function of perspiration.

Baublys (1975: 472/3) repeats the objection of Frankfurt and Poole (1966: 71) and reformulates it as follows:

> The set of belief-contravening suppositions required to evaluate [a functional counterfactual] involves the suspension of various laws of nature (Baublys 1975).

To clarify this issue let me first point out an unclarity in Canfield's account: should the normal organism be compared to a possibly real but non-normal organism or to a hypothetical one (that need not be physically possible)? Canfield is not clear on this point. His clarification in terms of action sentences suggests a comparison with a hypothetical and not necessarily physically possible organism. However in his example he compares a normal human with a non-normal human that does not exists but could be created experimentally by adding a pulse dampener to the heart of a normal human. By adding a pulse dampener to the description of an organism he does more than crossing out action-sentence chains in that description. It looks as if he quietly sacrifices his explicit requirement that the description of the comparative organism is obtained *only* by crossing out the activity under study for the implicit requirement that the comparative organism is physically possible.

If Canfield would require that the real organism is compared with a physically possible organism, then the objection of Frankfurt and Poole, that there can be no organism that is structurally the same as but functionally different from the organism under study, is obviously both relevant and tenable. In response, Canfield could try to modify his *ceteris paribus* clause to the effect that the comparative organism must be structurally as similar to the organism under study as is physical possible without performing the activity under study. This response would have several problems. First, there is no guarantee that this criterion uniquely determines a counterfactual situation. This is a problem if the survival value differs with respect to the different counterfactual situations. I will discuss this issue in the next section. Second, the structural changes needed to remove the activity under study might have side effects which decrease the survival value. As Frankfurt and Poole (1966) point out, it is far from obvious that a pulse dampener does not decrease the chances of survival of the organism concerned: “after all, it has an instrument added to its heart” (p. 72). This instrument and the modifications needed to attach it to the heart will have many effects apart from dampening the pulse and these effects could decrease the life chances of the organism that has the dampener. Nissen's counterexample is less convincing: in contrast to the plastic covering, the modification of the lungs is of no use in preventing the effect under study (transpiration) and, for that reason, the comparative organism used in producing the counterexample is less similar to the organism under study than is required by this modified *ceteris paribus* clause. Third, this modified clause makes functions dependent on our technical possibilities. If a certain effect is prevented with rude means that have many additional side effects, the chance that there is a negative influence on reproductive success is much larger than if the effect is prevented by more sophisticated instrument. But this means that if a more sophisticated instrument is developed, the life chances of the comparative organism could change and thereby the function of the activity under study. This is unacceptable.

On the other interpretation of Canfield's account the real organism must be compared with a hypothetical organism that has the same structure as the real organism but in which the activity under study does not take place. Of course, such an organisms cannot exist, but what does that matter? Canfield might maintain that he has given a recipe that allows one to determine the effects of the removal of the activity in study and that the fact that there can be no organism that satisfies the description used in that procedure is not relevant. He might add that experimental modifications of real organisms might yield information about the chances of survival of the hypothetical organism, but are of course not meant to create it (since this is not possible). This construal would answer both the non-uniqueness objection (Canfield's crossing out procedure yields a determinate description of all the activities the comparative organism may perform) and the objection put forward by Nissen (neither are causes crossed out nor are instruments added). Nevertheless, this line of defence does not work, as I will argue now. The main point of my argument is that it is not possible to construct a hypothetical organism if one does not require that this hypothetical organism satisfies the laws of nature.

I first argue that Canfield's recipe for determining the effect of the removal of a certain activity on the chances of survival by crossing out certain action sentence chains does not work and that in order to construct the hypothetical organism one has to calculate the effect of the removal in detail. On Canfield's construal all action-sentences have the form 'X does Y'. Canfield explicitly allows that a certain term in an action sentence chain (for example 'B' in <'A does B', 'B does C'> refers to an object (say bile salts) in one sentence and to an activity (for example the activity of bile salts) in the next. Furthermore the word 'does' is used for a great variety of activities, for example the first item in an action sentence can produce the second (as in 'folic acid synthesises methyl groups'), the first can secrete or excrete the second ('the liver secretes bile'), the first can cause or prevent the occurrence of the second ('heparin prevents blood clotting'), the first can take part in the second activity ('biotin participates in pyruvate metabolism'), and so on. This strategy of concatenating action-sentences into chains by means of common terms results in too many sentences crossed out.

Consider the following example. At least in mammals the production of hormones (such as the production of thyroxin (TH) by the thyroid) is often regulated by means of hormones (for example, the thyrotropic hormone (TSH) produced by the pituitary stimulates the thyroid to produce thyroxin). TH is also produced in absence of TSH but in lesser quantities. So, one finds a certain organ (the pituitary) that produces a hormone (TSH) which stimulates another organ (the thyroid gland) to produce another hormone (TH). Thus, we have the following action-sentence chain: <'the pituitary produces TSH', 'TSH stimulates the thyroid gland', 'the thyroid gland produces TH'>. Hence, if one wants to determine the function of TSH production one must cross out the sentence 'the thyroid gland produces TH' from the description of the organism. Nevertheless, a real organism without TSH will produce some TH. And how about the following chain: <'the thyroid produces TH', 'TH *inhibits* the hypothalamus', 'the hypothalamus produces TRH'>? According to Canfield's recipe, to determine the function of the production of TH one must cross out the production of TRH from the description of the organism. However, a real organism produces *more* TRH if TH is absent than if it is present. This example shows that one cannot construct a hypothetical organism simply by crossing out action-sentences. One has to calculate the effects of not performing a certain activity in detail.

In the paragraph above I argued that Canfield's recipe for constructing a hypothetical organism that does not perform the activity under study does not work as it should, and that in order to construct such a hypothetical organism one has to calculate the effects of not performing that activity in detail. This brings me to the main point of my argument. In order to make such a calculation one needs the laws of nature. So, if Canfield does not require that the hypothetical organism satisfies the laws of nature there is no way to calculate the way in which that hypothetical organism works.

It will be clear from the above account that I agree with Canfield's critics that counterfactual comparison does not work if one allows that the counterfactual organism does not satisfy the laws of nature. In my account counterfactual comparisons compare a real organism with a hypothetical organism that lacks the trait in study, or possesses it to a less degree. Yet, it is required that this hypothetical organism is physically possible though it need not be viable. Physically possible means that the hypothetical organism does conform to the laws of nature, but it need not be the case that it can be generated (experimentally or in nature).

6.3.5 How to determine the counterfactual situation

Many philosophers have objected to an analysis of 'function' in terms of counterfactual comparison that it is not clear with what counterfactual state of affairs the actual state is to be compared. For example, Frankfurt and Poole object against Canfield's (1964, 1965) analysis that the situation for comparison is not uniquely determined:

> There are an infinitude of possible mouse-like organism which do not secrete bile. Which of these does Canfield propose that we compare with the bile-producing specimen? (Frankfurt & Poole 1966: 71).

The idea that philosophers should provide a method that uniquely determines a counterfactual situation surfaces in Wimsatt's account too. Wimsatt (1972: 56/7) lists three desiderata which criteria that single out a reference situation should possess. One of them is that "they should pick out a unique situation for comparison", the other two are that "the criteria should be applicable to all of the functional traits", and that "the functional hierarchies picked out by the criteria should be highly similar to the functional hierarchy being analysed". He argues that none of the criteria available in biology satisfies these desiderata and suggests that an approach that combines the ideas of Simon (1957, chapter. 1) and Rescher (1964) is most promising.

The idea that there should be a method to determine a unique reference situation rests on the confusion of causal role and survival value prevalent in the survival value approach. If counterfactuals were used to single out causal roles there should be a unique reference situation. This is the case because the causal roles of an item depend on the way in which the organism works, but not on the structure of another organism with which it is compared. The human heart has the role to pump blood and this does not change if the human heart is compared to a frog heart or to an insect heart. So if causal roles were determined in comparison with a counterfactual situation one would need a uniquely determined reference situation. Otherwise the causal role could change if the reference situation changed. On the other hand, attributions of survival value are relative to the situation used for comparison. Gills have more survival value than lungs for an animal that extracts its oxygen from water, lungs have more survival value than gills for animals that breath air. Since attributions of survival value are essentially comparative the objection that counterfactuals do not pick out a unique reference state is not relevant. The fact that the survival value of a certain state or behaviour with respect to one counterfactual reference situation may differ from the survival value of that state or behaviour with respect to another counterfactual reference situation is not more problematic than the fact that the survival value with respect to one real reference situation may differ from the survival value with respect to another real reference situation. It just shows that attributions of survival value are essentially comparative and relative to a certain environment. Hence, the fact that there is no unique counterfactual situation for comparison does not count against counterfactual comparison.

More recently a related objection has been brought up by Millikan (1989a, 1993a) against the variant of the survival value approach defended by Bigelow and Pargetter (1987). Canfield and Wimsatt view functions as effects that increase the chances of survival of an organism in a sufficient number of cases. Bigelow and Pargetter emphasize that functions are of a dispositional nature: to count as a function an effect need not actually increase the survival of the organism, it is sufficient if it would enhance those chances in its natural environment.

> Something has a (biological) function just when it confers a survival enhancing propensity on a creature that possesses it (Bigelow & Pargetter 1987: 192).

Bigelow and Pargetter do not discuss the comparative nature of this “survival-enhancing propensity”. However, they note that propensity theories may vary in the way they work out this notion:

> [Propensity] 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 (Bigelow & Pargetter 1987: 194).

This quote suggests that Bigelow and Pargetter think of survival value in counterfactual terms: a trait has survival value if an organism that has this trait has a greater propensity to survive than an organism that would lack this trait.

Millikan objects that it is not clear how the counterfactual organisms is built, how it behaves and in which environment it lives:

> Unfortunately, exactly in this context, counterfactuals are notoriously indeterminate in truth value. If a given individual with a certain trait were *not* to have it, what would this individual have instead? There is no such thing, for example, as being simply *not monogamous*. Is the individual then to be celibate? Or homosexual? Or polygamous? If polygamous, how many wives does he juggle? How does he employ them? What do others in the community do about it? Are they monogamous, for example? Suppose that you didn’t have a nose. Well, would you have gills instead? Or maybe a trunk? Or just two holes? A closed flap over the two holes so that you must breath through your mouth? What would you do without eyes? Well, you might have radar in front, or bats’ ears. There is really *no sense at all* to the question how much, if at all, your monogamy, your nose or your eyes “enhance” your fitness, without first answering the question, *Enhances it over what?* (Millikan 1993a: 39/40, emphasis in original).

She adds to this:

> The notion of superior fitness, as actually used in evolutionary biology, is [...] understood relative to alternative traits *actually found in the population.* A moment’s reflection shows that this is indeed the only way to unpack the “enhanced” in “enhanced fitness” so as to lend it substance (Millikan 1993a: 40, emphasis hers).

This way of unpacking the notion of ‘enhancement’ is, of course, fatal to the survival value approach. For, on this construal, traits that do not vary in a population would not have a function. As Millikan puts it:

> On this reading [...] not only is it not the function of noses to support eyeglasses, but noses have no functions at all, unless the current population contains a portion of genetically noseless people who have, on average, fewer progeny that the rest of us (Millikan 1993a: 40).

Bigelow & Pargetter confuse causal role and survival value and for that reason I do not want to defend their analysis of ‘function’. However, their explication is on the right track (just as that of Canfield and Wimsatt) if it is taken as an explication of ‘survival value’. Millikan is wrong in thinking that the notion of ‘relative fitness’ (fitness as compared to variants actually present in the population) is the only sensible way to unpack the idea that a trait enhances the fitness of the organism that has it. As my examples in section 6.3.2 show functional biologists often unpack this idea in terms of counterfactual comparisons. As said before, counterfactual comparisons concern a hypothetical organism that is similar to the real organism except that the trait of which the survival value is studied is absent or present to a lesser degree. The comparison aims to establish not only what the survival value of a certain trait is but also under what conditions the trait under study has survival value. This means that at an initial stage of research the hypothetical organism is usually not well-defined. In the course of the study the description of the hypothetical organism and the conditions in which it lives become more and more refined. As the examples in section 6.3.2 show, counterfactual statements made by biologists often do not contain a detailed description of the counterfactual condition. However, when the context (chapter, paper, talk etc.) in which the statement is presented is taken into account it is usually clear enough with what situation the real situation is compared. In the context of a research paper it will be clear whether monogamy is compared with polygamy or celibacy and what the other organisms in the population do. If it is not clear that is a reason to blame the study concerned, not to reject the use of counterfactuals. This shows that Millikan’s argument is untenable as an argument against the use of counterfactual comparisons to determine survival value.

6.3.6 Are counterfactual judgements too complex to be evaluated?

Another complaint about counterfactuals is that they are too complex to be evaluated in a reliable way. This complaint is most clearly voiced by Baublys (1975) but it might be at the back of the mind of many others. Baublys maintains that the task to work out a description of an organism that does not circulate blood “would surely daunt even the most brave hearted and knowledgeable of physiologists” (p. 474). He adds that even if this problem were solved there remains the further problem of calculating the chances of survival of both the real and the non-normal organism. This makes counterfactual judgements very speculative and the fact that biologists agree about the functions of many parts of the body shows that they do not make such judgements:

> Biologists do not in practice disagree all that often in their functional ascriptions; and this surely supports the view that extremely speculative counterfactual comparisons do not play as basic an epistemological role as Wimsatt would have us believe (Baublys 1975: 476).

I would agree with Baublys that counterfactual comparisons are not used to single out causal roles. However, as I have shown in section 6.3.3 counterfactual comparisons do play a basic role in biology, namely in judgements of survival value. Philosophical analyses should account for this use rather than deny it. The fact that there is consensus among biologists about many functional counterfactuals shows that it is not as complex a task to work out a counterfactual organism as Baublys thinks. It also shows that determining the chances of survival of such a counterfactual organism is less speculative than Baublys thinks. Evaluation of counterfactual statements is often difficult but if the underlying physics is known there is hope of success. That is one of the reason why functional biology is such an interesting subject. Of course, there are many cases in which no decisive conclusion is reached but there are also many functional counterfactuals about which a verdict is given. Baublys remark that the task to work out a description of an organism that does not circulate blood would daunt even the most brave hearted and knowledgeable of physiologists is demonstrable false. The functional counterfactual that most of the organisms that have a circulatory system would not be able to diffuse oxygen if they lacked that system was conclusively established by Nobel prize winner August Krogh around 1940. It was not an easy task to establish this counterfactual but the difficulties were in the physics involved and in the instruments needed to measure the relevant quantities. Sure, it took him some work to figure out the conditions under which a circulatory system is needed, but it was not an impossible task and his success is uncontroversal. This shows that counterfactual comparisons are neither as speculative nor as complex as Baublys maintains they are.

6.3.7 Conclusion

Proponents of a survival value approach to function and functional explanation tend to explicate the notion of function in terms of counterfactual comparisons. I have shown, by many examples, that functional biologists are really involved in counterfactual comparisons and that, for that reason, no account of functional explanations should ignore them. I have also argued that in regard to counterfactual comparison one should distinguish between function as causal role and function as survival value. Counterfactual comparisons are used to determine the survival value of a certain trait but that they are not fit to single out causal roles. Both the proponents of the survival value approach and their critics tend to confuse causal roles and survival value. I have argued that the arguments brought up against the use of counterfactuals to determine function may hold against the use of counterfactual comparisons to determine causal roles (function_2), but do not count against the use of counterfactual comparisons to study survival value (function_3).

The conclusion that design explanations centre around functional comparisons is a strong argument in favour of my thesis that such explanations are not causal in character: hypothetical organisms that never existed do not affect the course of events. Hence, appeal to such organisms can not explain how a certain trait was brought about. This conclusion also gives rise to what I think is the main issue concerning functional explanation: what does one learn about a real organism by comparing it with hypothetical organisms? To put the point in a more philosophical style: how can it be explanatory to compare a real organism with a hypothetical one?

6.4 Adaptation

6.4.1 Introduction

Ruse (1973) proposes to analyze the notion of ‘function’ in terms of adaptation. He argues that no one would attribute to the long hairs of dogs the function to harbour flies unless harbouring flies contributes to the dogs ability to survive and reproduce (e.g. if flea bites would provide immunity to parasites), that is unless harbouring flies were an adaptation.

> The claim ‘the function of *x* in *z* is to do *y*’ implies that *y* is the sort of thing which aids the survival and reproduction of *z*. Now this is the kind of thing which [...] biologists call an ‘adaptation’ (Ruse 1973: 184).

On Ruse’s account functional statements of the form

> the function of *x* in *z* is to do *y*

should be analyzed as saying:

> (i) *z* does y by using *x*
> (ii) *y* is an adaptation

The latter statement (ii) is supposed to be equivalent to an attribution of survival value (ii"):

> (ii") *y* is the sort of thing which helps in survival and (particularly) reproduction (Ruse 1973: 186/7).

Unfortunately, as I will discuss, the term ‘adaptation’ as it is used by biologists is as ambiguous as the term ‘function’ and much more ambiguous than the term ‘survival value’. For that reason no clarity is gained by substituting ‘adaptation’ for ‘survival value’. However, Ruse is right that at least on the notion of adaptation most widely used by biologists, saying that a trait is an adaptation is more or less equivalent to saying that that trait has survival value. For that reason, discussions of the explanatory force of ‘adaptation’ might yield insights in the explanatory force of attributions of survival value. Indeed, it will turn out that there is at least one review on adaptation (Reeve & Sherman 1993) that is highly relevant to my topic.

The notion of ‘adaptation’ deserves a separate study and I will not attempt a complete review of the literature. Neither do I claim to provide new insights. I merely aim to pave the way to the paper of Hudson Reeve and Paul Sherman. The term ‘adaptation’ is used in at least five different ways: (1) adaptation as a phenotypic process, (2) adaptation as the fit between an organism and its way of life, (3) adaptation as a genotypic process, (4) adaptation as a fitness enhancing trait, and (5) adaptation as a trait built by natural selection. All these notions denote some way in which an organism is adjusted to its environment or an item to its role (function$_2$). The first and the third notion refer to processes, the others to properties of items or traits. The first notion refers to the processes by which an individual organism may become adjusted to its environment in the course of its life history. There are many such processes. Some are reversible others not. The term ‘adaptation’ is usually reserved for the reversible processes, such as the process by which organisms adjust their physiology in response to climatic changes or to changes in food quality. Well known cases are the tanning of the skin when it is exposed to the sun and the increase of the number of red corpuscles in the blood of humans who move to high altitudes. These processes and their differences are not relevant to my present purposes and I simply want to set them aside under the banner ‘adaptation as a phenotypic process’ (adaptation$_1$). The other concepts of adaptation are more pertinent to my purposes and I will discuss them in that order.

6.4.2 Adaptation as the fit between an organism and its way of life (adaptation$_2$)

The second notion of adaptation, adaptation as the fit between an organism and its way of life (adaptation$_2$), refers to the phenomenon that the way in which an organism is built and the way that organism behaves is particular well fit to a certain style of life. For example, the strong claws and keen sense of a carnivore are well made to catch preys and the parts of its alimentary canal are perfectly fit to digest flesh. The tail, beak and tongue of a woodpecker are apt to catch insects under the bark of trees, and so on. This phenomenon was, of course, already known to Aristotle and got a prominent place in biology around the turn of the eighteenth century, due to the work of both the French zoologist Cuvier and a group of German biologists involved in what is now called “the teleo-mechanic research program”. Cuvier is the founding father of the French tradition of morphology, one of the traditions that shaped biology in the nineteenth century (see Coleman 1964, Appel 1987). The teleo-mechanic research program was the leading research program in German biology during most of the nineteenth century and another source that shaped biology (see Lenoir 1982). Both Cuvier and the teleo-mechanists thought of functional morphology as the heart of a new science which they called ‘biology’. In their view one of the main aims of functional morphology is to explain the organization of animals by showing how the parts of those animals are adapted$_2$ to each other and to a certain way of life. The phenomenon of adaptation$_2$ itself, however, must remain unexplained.

As is well known the phenomenon of adaptation$_2$ also has a prominent role in Darwin's theory of evolution. Darwin probably took over this notion from English natural theology. Natural theology is best viewed as an attempt to link popularized science, theology and political conservatism. In Natural Theology the notion of perfect adaptation has a twofold explanatory role. First, it is used to phrase the phenomenon to be explained: the phenomenon of adaptation$_2$ is the phenomenon that the structure and behaviour of an organism fits a certain way of life. In Natural Theology this phenomenon is explained as the result of God's intentional design. God designed each kind of organism to perform a certain style of life. He made them in such way that each kind of organisms has the attributes that allow them to perform their way of life as good as possible. Because the ways of life of different organisms are adjusted to each other, the performance of each style adds to the glory of nature as a whole.[10] Second, the notion of adaptation$_2$ is used to explain the structure, activity and behaviour of an organism of a certain kind. This done by showing how these features adapt$_2$ the organism to its style of life. Again this makes sense on the assumption of intentional design: God gave that organism those features precisely because those features adapt$_2$ that organism to its style of life.

Darwin emphasizes that contrary to the presuppositions of natural theology the fit between the structure and behaviour of an organism and its way of life (adaptation$_2$) is not perfect. According to Darwin adaptation$_2$ is a matter of degree. He acknowledges that there are many structures which are almost perfectly adapted$_2$ to a certain way of life, but he explains this phenomenon in a novel way. On Darwin's theory the phenomenon of adaptation$_2$ (the phenomenon that the structure and behaviour of many organisms reaches a high degree of perfection in performing a certain style of life) is to be explained as the result of the process known as evolution by natural selection. The key to Darwin's explanation of adaptation$_2$ is the idea that differences in adaptedness$_2$ result in differences in reproductive success. As the result of small, heritable differences in structure and behaviour the organisms of a population differ in the degree to which they are adapted$_2$ to their style of life. Those organisms that are better adapted$_2$ to their style of life produce more offspring than those that are less well adapted$_2$. As a result the share of organisms that have traits that improve their adaptedness$_2$ increase in the next generation. A high degree of perfection of adaptation$_2$ results from the accumulation of innumerable of such small improvements in adaptedness$_2$ over many, many years.

In present-day evolutionary biology the explanatory force of the notion of adaptation$_2$ is highly contended. The notion of adaptation$_2$ has three possible explanatory roles. First, as in Natural Theology the phenomenon of adaptation$_2$ serves as a phenomenon to be explained. In Natural Theology the phenomenon that the structure and behaviour of every organism is per-

---

[10] The conservative moral is, of course, that in both the natural and the political order seeming imperfections increase the perfection of the whole.

fectly $adapted_2$ to the way of life of that organism is explained as the result of God's intentional design. On Darwin's theory the fact that many organisms are $adapted_2$ to a high degree to the style of life they have adopted is explained as the result of the accumulation of small improvements in $adaptedness_2$ due to natural selection. Second, differences in $adaptedness_2$ serve to explain differences in fitness. The better an organism fits its way of life, the better its chances to produce many offspring. This idea replaces the idea in Natural Theology that one can explain why an organism has the traits it has by showing that those trait $adapt_2$ the organism to its way of life. In Natural Theology, this idea makes sense on the assumption that God decided to give that organism the traits it has precisely because they $adapt_2$ that organism to the way of life he wants it to have. On Darwin's theory one can explain why the organisms of a certain generation have the traits they have by showing that in past generations the organisms having those traits were better $adapted_2$ to their way of life than their competitors that lacked those traits. This makes sense because on Darwin's theory the current organisms evolved the traits they have because those traits increased the $adaptedness_2$ of past ancestors. However, it is argued that on current formulations of the theory of evolution by natural selection the notion of $adaptation_2$ is superfluous. What is important in natural selection is the existence of heritable differences in fitness between variants in a population. These differences are to be explained on the basis of morphological, physiological and behavioural differences and there is no need for something like the fit between an organism and its way of life as an intervening factor. The geneticist Krimbas (1984) has gone as far as to argue that the concept of adaptation is not only superfluous but even detrimental to evolutionary theory and should therefore be excluded from scientific texts[11]. Third, many biologists attempt to explain the structure, activity and behaviour of an organism by showing how it $adapts_2$ that organism to its style of life. This attempt is known as "the adaptationist program". The program makes sense on the assumption that a good fit between the organism and a certain way of life shows that the organism has been built by natural selection to perform that way of life as good as possible. This assumption has been criticized by Stephen Gould and Richard Lewontin (1979), among others. Gould and Lewontin argue that a good fit between an organism and its way of life ($adaptation_2$) may result from other processes than selection to perform this way of life as efficient as possible. For example, a trait might have been co-opted for the task it currently performs after it evolved due to selection for another task (this kind of process is called 'change of function'). The use of the lung as a swim bladder in most teleost fish is a case in point. Processes like genetic drift and genetic linkage might have had a role too.

---

[11] Krimbas does not discriminate the different concepts of adaptation I have distinguished but he seems to be concerned with both adaptation as goodness of fit ($adaptation_2$) and adaptation as a trait that enhances fitness ($adaptation_4$).

> The mere existence of a good fit between organism and environment is insufficient evidence for inferring the action of natural selection (Gould & Lewontin 1979: 593).

Although the idea of a structure or behaviour fitting a certain style of life is intuitively clear in the face of such examples as the carnivore and the woodpecker it proved difficult to develop a notion of adaptation$_2$ precise enough to be applied in scientific research. The best attempt is that of Walter Bock and Gerd von Wahlert (1965) (see also Bock 1980). Bock and Von Wahlert observe that in order to stay alive an organism must spend energy in maintaining a bond with the environment. They also observe that at any given time an organism has only a limited amount of energy available. It is, therefore, advantageous for the organism to spend as little energy as possible in performing a certain task: the less energy an organism uses to perform its daily tasks, the more energy remains for that organism to meet unexpected or strenuous conditions. This justifies a definition of the degree of adaptation$_2$ as the inverse of the amount of energy needed to perform an essential task (essential to maintain the bond with the environment). For example, many birds are able to cling to vertical surfaces for some time but woodpeckers are better adapted$_2$ to this way of life than house sparrows because the amount of energy required for this activity is much lower for good climbers such as woodpeckers than for poor climbers such as the house sparrow.

The main problem with the definition of Bock & Von Wahlert is that energy economy is not always a good measure of the extent to which an organism fits a way of life. Consider for example the famous case of industrial melanism (Kettlewell 1973). Many species of moths vary in colour pattern: there are darker and lighter forms. In industrial areas the trees on which these moths rest during the day are covered with a layer of dark soot and for that reason in these areas dark moths are better camouflaged than light ones. However, on Bock & Von Wahlert's definition of adaptation$_2$ it would not be justified to say that the darker forms are better adapted$_2$ to life in industrial areas than the lighter forms because it does not cost less energy to be better camouflaged.[12]

The above considerations show that is not possible to give a more precise definition of adaptation$_2$ (more precise than something like 'the fit between an organism and its way of life' or 'how well a certain item fits its role') except in terms of evolutionary criteria such as fitness. Nevertheless in many cases it is possible to establish adaptation$_2$ without employing such evo-

---

[12] Brandon (1978) gives a more abstract argument. His argument is that it is always possible (at least in principle) to prevent those individuals that use less energy to perform a certain task from breeding while allowing the ones with higher energy requirements to breed. This shows that there is no lawlike relation between adaptation in Bock & Von Wahlert's sense and expected reproductive success (e.g. fitness). This lawlike relation is required if adaptation is to explain fitness. Brandon adds that a similar objection would apply to all definitions of adaptation that do not define adaptation in terms of fitness.

lutionary criteria. For example, we can judge how well eyes fit their roles on the basis of physical criteria. It turns out that the optical design of many eyes approaches optima predictable from physics (Goldsmith 1990).

6.4.3 Adaptation as a genotypic process (adaptation$_3$)

The third notion of adaptation, adaptation as a genotypic process (adaptation$_3$), refers to the process of genetic change that results in adaptation$_2$. This is one of two ways in which the word ‘adaptation’ is defined by Futuyma[13] in the main text of what has been the only good text on evolutionary biology for more than a decade:[14]

> Sometimes the word [‘adaptation’] refers to the process whereby a population is altered in such way as to be better suited to its environment (Futuyma 1986: 251).

Defined in this way any process of genetic change that results in adaptation$_2$ would count as a process of adaptation$_3$. Many authors restrict adaptation$_3$ to the process of natural selection. This is what Futuyma does in his glossary, where he defines ‘adaptation’ as

> a process of genetic change of a population, owing to natural selection, whereby the average state of a character becomes improved with reference to a specific function, or whereby a population is thought to have become better suited to some feature of its environment (Futuyma 1986: 550).

The same kind of definition can be found in Kluge’s well-known textbook on functional morphology:

> *Adaptation* can be defined as the hereditary adjustment of an organism to its environment by means of natural selection (Kluge 1977: 7).

6.4.4 Adaptation as a fitness enhancing trait (adaptation$_4$)

The problems in making the notion of adaptation as the fit between an organism and its way of life (adaptation$_2$) more precise have led many biologists to replace this notion by that of adaptation as a trait that enhances fitness (adaptation$_4$). Nowadays, most evolutionary biologists define adaptation in this way. This kind of definition is, for example, adopted by Keeton and Gould in their famous introduction to biology:

> In biology, an adaptation is any genetically controlled characteristic that increases an organism’s fitness. Fitness, as the term is used in evolutionary biology, is an individual’s (or allele’s or genotype’s) probable

---

[13] The other is that of adaptation$_5$.

[14] The first impression of Futuyma’s book appeared in 1979. Until the publication of Ridley’s (1993) book on evolutionary biology there was no alternative to Futuyma’s book.

> genetic contribution to succeeding generations. An adaptation, then, is a characteristic that enhances an organism's chance of perpetuating its genes, usually by leaving descendants (Keeton & Gould 1993: 473).

and also by John Alcock in his well-known text on behavioural biology:

> We shall define an adaptation as an inheritable characteristic that gives an individual an advantage over others with different inherited abilities, an advantage in transmitting its genes to subsequent generations. An adaptation is *better* than other alternatives that exist, better than it would be if it were slightly modified, better at "helping" individuals pass on their genes (Alcock 1989: 218).

In the glossary of Mark Ridley's textbook on evolutionary biology 'adaptation' is defined along the same lines:

> Feature of an organism enabling it to survive and reproduce in its natural environment better than if it lacked the feature (Ridley 1993: 631).

In the main text Ridley acknowledges that there are two ways to define adaptation; one defines an adaptation as "any character which helps its bearer to survive and reproduce", the other limits the application of the term "adaptation" to "organs that are still serving the function they originally evolved to do" (p. 331). These two ways correspond to my notion of adaptation as a fitness enhancing trait (adaptation$_4$) respectively adaptation as a trait built by natural selection (adaptation$_5$). Ridley thinks that the first definition (adaptation$_4$) is "probably the one more widely accepted among biologists who are actively doing research on adaptation" (p. 331).

The definition of adaptation as a fitness enhancing effect (adaptation$_4$) is also adopted by at least one of the founders of the modern synthetic theory of evolution, namely Theodosius Dobzhansky:

> An adaptive trait is [...] an aspect of the developmental pattern which facilitates the survival and/or reproduction of its carrier in a certain succession of environments (Dobzhansky 1956: 347)

He repeats this in 1968:

> An adaptive trait is structural or functional characteristic, or more generally, an aspect of the developmental pattern of the organism, which enables or enhances the probability of this organism surviving or reproducing (Dobzhansky 1968: 6/7).

As I said at the beginning of this section, the notion of adaptation as a trait enhancing fitness (adaptation$_4$) was meant to replace the notion of adaptation as fit between the organism and its way of life (adaptation$_2$) in Darwinian biology. However, the notion of adaptation as a trait enhancing fitness (adaptation$_4$) cannot fulfil the second explanatory role attributed to the notion of adaptation as fit between an organism and its way of life (adaptation$_2$) in Darwinian biology, namely to explain fitness. The reason is that on this notion saying that a trait is an adaptation$_4$ is the same as saying that it enhances fitness, rather than explaining that it does so. It has been debated whether or not the notion of adaptation$_4$ can perform the two other explanatory roles attributed to adaptation$_2$ in Darwinian biology, namely explicating a phenomenon to be explained by natural selection (as an alternative to divine design) and helping to explain why the organism is built the way it is built. Reeve and Sherman (1993) provide a definition of adaptation$_4$ that according to them does both:

> An adaptation is a phenotypic variant that results in the highest fitness among a specified set of variants in a given environment (Reeve & Sherman 1993: 9).

On this notion the phenomenon to explained by appeal to natural selection is the phenomenon that among a specified set of phenotypes the variant that is the most adapted$_4$ one in a specified environment is the one that in fact prevails in that environment. Note, that the statement that a certain phenotypic variant is both the one best adapted$_4$ and the prevalent one among a specified set of alternatives in a certain environment is a testable hypothesis, that is neither true by definition, nor by definition due to natural selection. Furthermore, if this hypothesis applies (that is, if the most adapted$_4$ phenotype is in fact the one prevalent among a specified set of alternatives in a certain environment), the best explanation we have is that it is maintained by natural selection. For that reason, if the most adapted$_4$ variant is the one most frequently found this is strong evidence for selection.

> Evolutionary biologists test selective hypotheses by determining whether the most common trait is really the one that maximizes some aspect of fitness relative to its alternatives (Reeve & Sherman 1993: 14).

To defend their analysis against the criticism (such as that of Gould & Lewontin 1979) that the fact that an organism fits its environment is not sufficient evidence for selection, Reeve and Sherman make use of the distinction between studies of evolutionary history and studies of phenotype existence (see Antonovics 1987, Brooks & McLennan 1991). The former studies are concerned with the modification of traits in the course of the history, the latter with their persistence in the current population. Reeve and Sherman agree with the critics of the adaptationist program that the fact that the prevailing trait is also the one most adapted$_4$ does not necessary imply that that trait became prevalent due to selection for its current use. Other evolutionary mechanisms such as genetic drift and changes of function might have played an important role too. However, they argue that this argument is irrelevant to their use of adaptation$_4$ because an appeal to the prevalence of the better adapted$_4$ variant is not meant as evidence that the trait has a selection *history* but as evidence that the trait is currently *maintained* by selection. The notion of adaptation as a trait that enhances fitness (adaptation$_4$) is of use in studies of phenotype existence. Studies of phenotype existence seek to explain the prevalence of a certain trait in a population by showing how that trait is maintained in the population, no matter the specific historical pathways leading to that prevalence. This is possible because genetic changes from one generation to another are determined by the characteristics of the changing generation and the environment in which it lives. The point is not that history is unimportant but that whatever is important about history must be laid down in the environment or in the organisms concerned, otherwise it would be ineffective. If it can be shown that the variant most frequently found in a certain environment is also the one most adapted$_4$ this is strong evidence that the trait is maintained by natural selection. For, according to Reeve and Sherman, natural selection is the only mechanism able to explain why an adapted$_4$ variant is not replaced by plausible but less adapted$_4$ alternatives. Alternative mechanisms such as lack of genetic variation, genetic linking, recurrent immigration and genetic drift are important to explain those cases in which the prevalent variant is not the most adaptive$_4$ one, but natural selection is the only mechanism that explains why a prevalent trait is adaptive$_4$.

How does the notion of adaptation as a trait that enhances fitness (adaptation$_4$) relate to the notion of adaptation as fit between the organism and its style of life (adaptation$_2$)? The notion of adaptation$_2$ in both Darwinian and pre-Darwinian biology primarily served to phrase a phenomenon to be explained (namely the fit between an organism and its lifestyle) and Darwin claims that his theory explains this phenomenon better than the creationist alternative. Reeve and Sherman show that this phenomenon can also be phrased in terms of the notion of adaptation$_4$, namely as the phenomenon that the variant that prevails in a certain environment is an adaptation$_4$, that is the one that has the highest fitness in that environment among a set of plausible alternatives. In Natural Theology and in Darwinian biology the notion of adaptation as the fit between the organism and its style of life (adaptation$_2$) also served to explain the way in which an organism is built and behaves. In Natural Theology this makes sense on the assumption that God made each kind of organism to perform a certain way of life and that he gave the individuals the attributes that allowed them to perform this way of life as good as possible. In Natural Theology showing that a certain trait adapts$_2$ the organisms having it to the style of life bestowed on those organisms is explanatory since God gave it that trait because it adapts$_2$ the organism to its style of life. In Darwin’s theory showing that in the past a certain variant was better adapted$_2$ than its rivals is explanatory since that variant was selected because it was better adapted$_2$ to its way of life. Reeve and Sherman argue that the notion of adaptation$_4$ can perform a similar role in Darwinian biology: if the prevalent variant is an adaptation$_4$ this is strong evidence that that variant is maintained in the population by natural selection.

6.4.5 Adaptation as a trait built by selection (adaptation$_5$)

The notion of adaptation as a fitness enhancing trait (adaptation$_4$) goes back to the times of the evolutionary synthesis (1940-1950). The last two decades another notion of adaptation, adaptation as a trait built by natural selection (adaptation$_5$), has gained users especially among biologists engaged in phylogenetic analysis (e.g. Brooks & McLennan 1991, Harvey & Pagel 1991). This notion of adaptation originates from George Williams’s famous *Adaptation and*

Natural Selection (1966). In the table of contents Williams declares that an effect of an item is to be called a function only if the item was designed to produce that effect (that is if the item evolved as a means to that effect — see section 2.2.4):

> Evolutionary adaptation is a special and onerous concept that should not be used unnecessarily, and an effect should not be called a function unless it is clearly produced by design, and not by change (Williams 1966: vii).

Similarly, to show that an item is an adaptation it is not sufficient to show that it has some beneficial effects, one must show that it has been designed to produce a certain effect.

> The demonstration of effects, good or bad, proves nothing. To prove adaptation one must prove functional design (Williams 1966: 212).

In the final chapter this is emphasized again:

> One should never imply that an effect is a function unless he can show that it is produced by design and not by happenstance. The mere fact of the effect's being beneficial from one or another point of view should not be taken as evidence of adaptation. [...] Parsimony demands that an effect be called a function only when chance can be ruled out as a possible explanation. In an individual organism an effect should be assumed to be the result of physical laws only, or perhaps the fortuitous effect of some unrelated adaptation, unless there is clear evidence that it is produced by mechanisms designed to produce it. (Williams 1966: 261).

Williams defines neither ‘function’ nor ‘adaptation’ explicitly but the idea is that an item is to be called an adaptation only if we have evidence that it is designed for some purpose. Evidence for design exists if the item is “too complexily organized” to be built by chance (e.g. mutation and genetic drift). Because the only known mechanism for natural design is selection this means that an item is to be called an adaptation only if we have evidence that it is built by natural selection.

Williams’s concept of adaptation$_5$ is in need of elaboration. One reason is that it remains vague what exactly is meant by the phrase ‘built by natural selection’. According to Williams the evidence to show that an item is built by selection is provided by showing that item is too complexily organized to be built by chance. But how complex is too complex? These problems are solved by Sober’s (1984) definition:

> *A* is an adaptation for task *T* in population *P* if and only if *A* became prevalent in *P* because there was selection for *A*, where the selective advantage of *A* was due to the fact that *A* helped to perform task *T* (Sober 1984: 208)).

Sober provides a clear explication of the phrase ‘built by selection’ (the trait must have become prevalent due to selection) and a clear account of what it means to be an adaptation for something. Note that on his definition adaptations are traits (not necessarily items) and adaptations are adaptations for a certain task (not for environments).

Sober's definition in his introductory *Philosophy of Biology* appears more permissive:

> Characteristic *c* is an adaptation for doing task *t* in a population if and only if members of the population now have *c* because, ancestrally, there was selection for having *c* and *c* conferred a fitness advantage because it performed task t (Sober 1993: 84).

On this definition it is not required that the trait spread through the population due to selection, it is even not required that the trait is prevalent it is enough if it was maintained by selection.

A definition similar to that of Sober (1984) is adopted by Futuyma in the glossary

> *An* adaptation: a feature that has become prevalent in a population because of a selective advantage owing to its provision of an improvement in some function (Futuyma 1986: 550).

The definition in the main text confuses adaptation$_4$ (adaptation as a trait that enhances fitness) and adaptation$_5$ (adaptation as a trait built by natural selection):

> An adaptation is a feature that, because it increases fitness, has been shaped by specific forces of natural selection acting on genetic variation (Futuyma 1986: 251).

The explanatory force of the notion of adaptation as a trait built by selection (adaptation$_5$) is different from the original explanatory role of the notion of adaptation in both pre-Darwinian and Darwinian biology (adaptation$_2$). As discussed in section 6.5.2, the notion of adaptation$_2$ served (i) to phrase a phenomenon to be explained (namely the fit between an organism and its lifestyle) and (ii) to explain the structure and behaviour of organisms. The notion of adaptation$_5$ is used to explain the structure and behaviour of organisms. More specifically it is used to explain how a certain trait spread through the population and acquired its current frequency. To say that a trait is an adaptation$_5$ for some achievement is a short way of saying that that trait spread through the population due to selection for an effect produced by that trait. This in turn is a metaphorical way of saying that that trait spread through the population because it had a certain effect that caused an increase of the relative fitness of the organisms having that trait. It makes sense to say that a trait spread through the population because variants having that trait were better adapted$_2$ than variants that lacked that trait. However, it does not make sense to say that a trait spread through the population because variants that had that treat were better adapted$_5$ than variants that lacked that trait.

Several authors (e.g. Coddington 1988, Brooks & McLennan 1991, Harvey & Pagel 1991) have emphasized that in order to determine what a trait is adapted$_5$ for one needs detailed information about its phylogenetic history. Because of the possibility of a change of function, to show that a trait has survival value compared to plausible alternatives (adaptation$_4$) due to a certain effect is not sufficient to show that the trait was built by selection for that effect. For example, in polar regions the white coat of a polar bear has survival value in comparison to a darker coat because it camouflages the bear. This does not show that the coat evolved as an adaptation$_5$ for camouflage, it might also have evolved for, say, thermoregulation. Furthermore, correlation between the presence of a trait among current species and their way of life does not show that that trait is an adaptation$_5$ to that way of life. The reason is that several different species might have inherited a trait from one species in which it originally evolved. In other words, one needs correlation between origins and ways of life, not between species and ways of life.

6.4.6 Conclusion

The term ‘adaptation’ is used in a number of different ways. For my purposes the interesting ones are that of adaptation as the fit between the organism and its way of life (adaptation$_2$), adaptation as a trait enhancing fitness (adaptation$_4$) and adaptation as a trait built by natural selection (adaptation$_5$). These different notions of adaptation have different explanatory roles.

Several authors have proposed to distinguish between different uses of ‘adaptation’ by taking advantage of the different grammatical forms (such as ‘adaptation’, ‘adaptive’, ‘adapted’, and ‘adaptedness’), others have coined new terms to deal with the distinctions. For example, Brandon (1981) proposes to restrict the application of the term ‘adaptation’ to the process of natural selection (adaptation$_3$) and to traits built by that process (adaptation$_5$). Brandon acknowledges that there are authors that have held the view that the term ‘adaptation’ refers to traits that are beneficial to their bearers (adaptation$_4$) but according to him this “minority view” (sic!) is “without any interest” since it “divorces adaptation from the evolutionary process”. Gould and Vrba (1982) coin a new term, ‘aptation’, for traits that enhance fitness (adaptation$_4$). They propose to restrict the term ‘adaptation’ to traits built by natural selection for their current use (traits that are adaptations$_5$ for the same role they are adapted$_4$ for) and coin the term ‘exaptation’ for traits that are currently apt (adapted$_4$) to perform a certain role but were not built by selection to perform that role. Endler (1986) proposes to restrict the term ‘adaptation’ to the genotypic process (adaptation$_3$) and to use the term ‘adaptive trait’ for a trait that enhances the fitness of an organism (adaptation$_4$).

The notion of adaptation$_4$ and especially Reeve and Sherman’s construal of that notion is akin to my notion of survival value. It will be clear that if a trait is an adaptation in relation to a certain set of variants on the definition of adaptation as a fitness enhancing trait (adaptation$_4$) it has survival value as compared to the traits of the other variants in the set on my definition of survival value. However, the reverse is not necessarily the case. The notions of ‘survival value’ and ‘adaptation as a fitness enhancing trait’ are both comparative. In the latter case (adaptation$_4$) the actual organism is compared with a set of *plausible* alternatives, that is with variants that could easily arise from the variants currently present in the population. These are the variants against with the real variant is actually maintained. The notion of survival value on the other hand is broader, it allows for comparison of the trait under study with traits that can not easily arise from the current population, such as hypothetical traits and traits present in other species.

Part II: The explanatory role of attributions of survival value

6.5 Canfield's account of the explanatory role of appeals to survival value

6.5.1 Canfield's account of functional explanation

According to Canfield (1964) statements of the form 'the function of *i* is *f*' are directly explanatory:

> Someone might say, 'Explain the function of the thymus', or ask, 'What is the function of the thymus?' or 'Why do animals have a thymus?' When we answer 'the function of the thymus is [such and such]' we have, it seems plain, given an explanation (Canfield 1964: 293).

On Canfield's account, statements of this form tell us what a certain item does that contributes to the capacity to survive and reproduce of the organisms that have it (see section 6.2.2 above). Canfield argues that the covering law model does not adequately account for the explanatory force of such attributions. According to Canfield the position that functional explanations conform to the covering law model has been stated most clearly by Arthur Pap (1962). According to Pap teleological explanations are characterized by the fact that

> a certain process in, or a certain characteristic of, an organism is explained as one that serves a certain purpose, as either a necessary or a sufficient condition for the achievement of a goal (Pap 1962: 359)

Consider a "very simple example" (p. 360) of such a teleological explanation:

> The heart beats in order to circulate the blood, which circulation in turn is necessary for the organism's survival (Pap 1962: 360).

According to Pap

> The heart's activity is here explained in terms of the function it serves. [...] The organism cannot survive unless the blood circulates and the blood cannot circulate unless the heart beats: that's why the heart beats (Pap 1962: 360).

This explanation fits the covering law model:

> To assert that the beating of the heart is a necessary condition for blood circulation is equivalent to asserting that blood circulation is a sufficient condition for the beating of the heart. And this is to assert the

> confirmable lawlike generalization that if blood circulates in an organism, then the organism contains a beating heart. Add the verifiable premise that blood circulates in this organism and you get deductively the conclusion that this organism contains a beating heart (Pap 1962: 360).

Hence, on Pap's view function attributions serve as explanations in accordance with the deductive-nomological model of the presence of the trait to be explained if the original function attribution is unpacked as the confirmable lawlike generalization and supplemented with a second premise stating initial conditions. In the case of the heart we have:

(1) The blood circulates only if the heart beats
(2) The blood circulates in Fido
----
(3) Therefore, Fido contains a beating heart (Canfield 1964: 294).

The first premise is a lawlike generalization, the second states initial conditions and the conclusion follows logically form the premises.

As I said in chapter 1, Pap (1962) and Nagel (1977) distinguish functional explanations from causal ones. According to these man the differences recede primarily in the kind of law appealed to. Functional explanations cite laws that mention consequences of the presence of the item the presence of which is to be explained, whereas causal explanations cite laws that mention the causes of that presence.

Canfield argues against this view that functional explanations differ from explanations that fit the covering law model both in their structure and in the questions they address. This is best seen when one compares the question a function attribution is meant to answer ('why does the heart beat?'), and the question addressed by Pap's reconstruction ('why does Fido contain a beating heart?'):

> This view of [functional explanation][15] is wrong. This becomes clear when we notice the disparity between the question which originally provokes a [functional explanation], and the answer given by the above set of premises. The question is: 'Why does the heart beat?' (Note Pap's 'That's why the heart beats'.) Whereas the above set of premises answers a different question, namely, 'Why does Fido (this organism) contain a beating heart?' (Canfield 1964: 294).

More generally, explanations that conform to the covering law model are not fit to answer the kind of questions biologists have in mind when they give functional explanations. Those biologists are interested in the question how a certain item is useful to the organisms that have it, whereas explanations that fit the covering law model explain the presence of a certain item.

---

[15] Canfield uses the term 'teleological explanation'.

> Explanations which fit the covering law model [...] attempt to *account* for something's being present, or having occurred by subsuming it under a general law, and by citing appropriate 'antecedent conditions'. [Functional explanations][16] in biology, of the kind we have been considering, do no such thing. They merely state what the thing in question does that is useful to the organisms that have it (Canfield 1964: 295).

Hence, according to Canfield there are two kind of differences between functional explanations and explanations that fit the covering law model. First, they differ in structure: functional explanations consists of a single function attribution, whereas explanations that fit the covering law model consist of premises stating covering laws, premises stating initial conditions, and a conclusion that follows logically from the premises. Second, they differ in the kind of question they address: functional explanations address the question 'how is this item useful to the organisms that have it?' whereas explanations that fit the covering law model address the question 'why is this item present in such and such organisms?'.

As I showed in chapter 4, Canfield is right that the explanations which biologists call 'functional explanations' (and which I have called 'design explanation') do not conform to the covering law model. However, as I show in the next section, Canfield's own account of functional explanation is unsatisfactory, as an account of design explanation. Canfield is right that a function attribution explains what an item is good for. Functional *explanations* (design explanations), however, do not consist of a single function attribution. They are much more complex and they explain much more than "what the thing in question does that is useful to the organisms that have it".

6.5.2 Function attributions and design explanations

Introduction

In the previous section I stated that Canfield maintains (i) that function explanations consist of a single function attribution, and (ii) that functional explanations tell us merely what a certain item is good for. On his view, the statement 'the function of the thymus is to initiate the differentiation of T-lymphocytes' explains why (certain) animals have a thymus and it does so by specifying what the thymus does that is useful to the animals that have a thymus. Similarly, the statement "the heart beats in order to circulate blood" explains why the heart beats and it does so by specifying what the heart does that is useful to those organisms that have a heart. In this section I aim to show, by means of examples, (i) that the kind of explanations which biologists call "functional explanations" and which I have called "design explanations are much more

---

[16] Canfield uses the term 'teleological explanations'.

complex than Canfield holds, and (ii) that Canfield's account of functional explanations as explanations that tell us what an item is good for leaves much out of sight of what is achieved by a design explanation. Design explanations address the much wider questions 'why is it useful that a certain item or behaviour has a certain character' or `why is it useful that a certain organism performs a certain activity'.

In support of my two theses I discuss a number of examples. Because of Canfield's confusion of attributions of causal roles and attributions of survival value I will examine the explanatory role of both of these attributions. I start with attributions of causal roles

Krogh's (1941) explanation of the circulatory system.

In section 4.2.4 I discussed explanations that explain why certain organisms have an item that performs a certain role by appeal to a need satisfied by the performance of that role. An example is Krogh's (1941) explanation of why many organisms (among others vertebrates) have a circulatory system. The train of thought in this explanation can be expressed as follows:

(1) Vertebrates are organism in which the distance between some organs and the periphery is more than one millimetre

(2) Organism in which the distance between some organs and the periphery is more than one millimetre, are able to survive only if they have a system of convection in addition to diffusion

(3) The circulatory system of vertebrates provides a system of convection in addition to diffusion

----

(4) That's why vertebrates have a circulatory system

This structure does not fit the covering law model, but it does not fit Canfield's model either: it does not consist of a single function attribution. The function$_2$ attribution (attribution of a causal role) (3) is combined with a statement pointing out that vertebrates are built in a certain way (1) and a lawlike statement (2) which states that a certain need arises in organisms that are built in the way specified in (1).

Furthermore, the question Krogh addresses is not Canfield's 'what does the circulatory system do that it useful to the organism to have?'. After all, the causal role of the blood in transporting oxygen was already known in the eighteenth century. Rather, Krogh addresses the wider question 'why is it useful to have a circulatory system?' In answer to this question he points to the way in which those organisms that have a circulatory system are built and shows with help of a law of physical chemistry that given the way those organisms are built they could not survive without a circulatory system.

Schwenk's (1994) explanation of why snakes have forked tongues

In section 4.3.3 I discussed explanations that explain the character of an item by appeal to its causal role (function$_2$). An example is Schwenk's (1994) functional explanation of why snakes have forked tongues. The train of thought in this explanation can be represented as follows:

(1) The tongues of snakes have a role in chemosensory tropotaxis
(2) Chemosensory tropotaxis is physical possible only if an organism is able to sense simultaneously the chemical stimuli at two points
(3) In snakes this requirement is met by the forking
----
(4) That's why the tongues of snakes are forked

Again, this structure does not fit the covering law model, but it does not fit Canfield's account either: it does not consist of a single function attribution and it does address a much wider question than 'what do forked tongues do that is useful for the organism to have?'. As the title of his paper indicates Schwenk addresses the question 'why are the tongues of snakes forked?'. The attribution of a causal role (function$_2$) is the first step in this explanation. Taken in isolation it explains what the tongue does that is useful for the organism to have. However, taken in isolation that attribution does not count as an explanation of why the tongues of snakes are forked. It is explanatory in this sense only because it is combined with a statement relating the causal role to a need (2), and a statement relating the character to be explained to that need (3). Schwenk discovered the trail following role of the snake's tongue. But he did much more. He uses this insight to explain the forked character of that tongue by pointing out that this character satisfies the need imposed on the tongue by that causal role, namely the need to sample chemicals at two points at one time. Canfield's theory of functional explanation fails to account for this second part of the explanation.

Habibi c.s. (1993) on gazelles.

In section 2.2.3 and 6.2.4 I discussed the comparison of the behaviour of sand gazelles and mountain gazelles and the explanation of the differences in terms of survival value by Habibi's and his colleagues (1993). Let us see how Canfield's account of functional explanations fares in the face of this study. First, consider the explanation of why mountain gazelles mark their territory with dung piles rather than with scent marks. The authors suggest that mountain gazelles do not use scent-marks because "scent-marks may be less long-lasting, and it would not be possible for a male to replenish them fast enough" (p. 51). The train of thought in this explanation can be represented as follows:

(1) Scent marks and dung piles are means to mark a territory
(2) An organism that marks its territory must be able to maintain the marking

(3) Mountain gazelles have large territories

(4) Scent marks do not last long

(5) Because of (3) and (4) a mountain gazelle would not be able to replenish scent marks fast enough to maintain the marking

(6) Dung piles last much longer than scent marks

(7) Because of (6), despite (3) a mountain gazelle is able to maintain dung pile marking

----

(8) That’s why mountain gazelles mark their territory with dung piles rather than with scent marks.

The structure of this explanation is basically the same as the structure of Schwenk’s explanation of the snake’s forked tongue. It starts by attributing a causal role (function$_2$) to the behaviour (in Schwenk’s case: the item) the character of which is to be explained (1), next it points to a need related to that causal role (2), finally it points out that if the character to be explained (dung piles) satisfies that need, whereas the alternate character (scent marks) does not (3-7). Accordingly, in this case Canfield’s account fails for the same reason as in the case of the snake’s tongue. First, the attribution of causal role to a certain behaviour is the first step in the explanation, rather than the explanation. Second, the explanation explains not only how the behaviour is useful but also why it is.

The explanation of why male sand gazelles herd females during the rutting season and male mountain gazelles keep large territories during the whole year has the following train of thought:

(1) Male territory behaviour in gazelles has a role in finding mates

(2a) If an animal breeds during the whole year at any time of the year the chance that a particular female will soon become receptive is fairly low

(3a) Because of 2a: if an animal breeds during the whole year it is more useful for the male to keep large territories during the whole year than to herd all females that come across his path

(2b) If an animal breeds seasonally, during the rutting season, a large proportion of females will be either receptive or about to become receptive

(3b) Because of 2b: if an animal breeds seasonally, it is more useful for the male to herd females during the rutting season than to keep large territories during the whole year

(4a) Mountain gazelles breed during the whole year,

(4b) Sand gazelles are seasonal breeders

----

(5) That’s why male sand gazelles herd females during the rutting season and male mountain gazelles keep large territories during the whole year.

This explanation relates differences in territory behaviour (large territories vs. herding) to differences in life style (breed during the whole year vs. seasonal breeding) in terms of survival value. Contrary to what Canfield maintains this explanation does not consist of a single function attribution. It starts with the attribution of the causal role to find mates to territory behaviour (1). Next it states that given a certain lifestyle (breed during the whole year) that causal role is better performed by one kind of territory behaviour (large territories) than by the other (herding females) (2a, 3a). Given another lifestyle (seasonal breeding) it is the other way round (2b, 3b). Because mountain gazelles have the first life style and sand gazelles have the second (4a,b) different kind of territory behaviours are appropriate to them.

Tinbergen c.s. (1962) on egg shell removal

In section 2.2.3 and 6.2.5 I discussed the study of the survival value of the egg shell removal behaviour of black headed gulls by Tinbergen and his colleagues (1962). Tinbergen and his colleagues carefully avoid to say that they explain that behaviour. Yet, it appears that their study reveals at least the beginnings of an explanation. The train of thought involved in this explanation runs as follows:

(1) The eggs of black headed gulls are subject to predation by herring gulls and carrion crows.
(2) Herring gulls and carrion crows find those eggs better if there is an empty egg shell in the proximity of the nest
(3) The removal behaviour prevents the empty egg shell from laying in the proximity of the nest
----
(4) That’s why herring gulls remove the empty egg shell after the chick has hedged

This explanation points to an effect of the removal behaviour (3) and shows that this effect has survival value given the conditions in which black headed gulls live (their eggs are subject to predation (1) by predators that find eggs better if there lays an empty egg shell nearby).

Conclusion

Finally, consider Canfield’s example of the thymus. Canfield suggests that the statement “the function of the thymus is to initiate the differentiation of T-lymphocytes” explains why “animals” have a thymus (I have already given this quote in section 6.5.1):

> Someone might say, ‘Explain the function of the thymus’, or ask, ‘What is the function of the thymus?’ or ‘Why do animals have a thymus?’ When we answer ‘the function of the thymus is [such and such]’ we have, it seems plain, given an explanation (Canfield 1964: 293).

It is true that the main insight about the thymus gained by the study of the thymus in the beginning of the 1960s is the insight that the thymus of mammals and birds has a causal role in the initial differentiation of T-lymphocytes. However, when compared to the examples above it is doubtful whether that attribution of a causal role (function$_2$) suffices as a functional explanation.

In his monograph on *Concepts and Approaches in Animal Morphology* the functional morphologist Peter Dullemeijer discusses the following attribution a causal role (function$_2$): “aquatic vertebrates have fins to move or to propel”. He maintains that this sentence is unsatisfactory as an explanation of why aquatic vertebrates have fins. The reason is that this attribution does not give us insight in the relation between fins and propulsion, that is in the relation between the form of the locomotory organs (they take the form of fins) and their causal role (the organism moves itself actively through water):

> The simple statement on aquatic vertebrates [“aquatic vertebrates have fins to move or to propel”] is unsatisfactory [..]. This discontent is not felt because of its simplicity, but because of a shortage of information to gain an insight into the *relation* between propulsion and fins (Dullemeijer 1974: 53, emphasis in original).

In order to explain why aquatic vertebrates have fins it is not enough to point out the causal role of fins, one must also explain why fins are useful to perform that causal role: why don’t aquatic vertebrates propagate themselves by lateral undulation of their body without using fins? In a similar vein McNeill Alexander requires that functional explanations show that the character of the item or behaviour in study is optimal to its causal role:

> We do not think a functional explanation complete until we can show that a structure or movement is optimal (by some plausible criterion) for the proposed function (McNeill Alexander 1988: 237)

The same considerations apply to the case of the thymus. The discovery that the thymus of mammals and birds has a causal role in the differentiation of T-lymphocytes is explanatory only in the sense that it tells us how the thymus contributes to the organism’s physiology. But that knowledge alone does not count as a functional explanation of why those animals have a thymus. It leaves unexplained why those animals have a special organ to perform that causal role. The first attempts to answer that question appeared in the late 1980s. Canfield is, therefore, wrong in identifying functional explanations with function attributions. An attribution of a causal role shows how the item to which the causal role is attributed contributes to the organism’s physiology. Functional explanations on the other hand explain the character or presence of an item or behaviour by appeal to their causal role and/or survival value. They often start with an attribution of a causal role, but taken in isolation this attribution is explanatory only to a limited extent.

My examples show that functional explanations do not consist of a single function attribution. They have a complex structure and involve a number of statements of different kinds. The most important ones are: attributions of causal roles, attributions of needs and/or survival value, statements specifying the conditions in which a certain item or activity is useful, and statements specifying what counts as being useful. An attribution of a causal role is often a first step in the explanation.

6.5.3 Conclusion

In the previous section I showed that Canfield’s account of functional explanation is unsatisfactory for three reasons: (i) it confuses attributions of causal roles and attributions of survival value, (ii) it misrepresent the structure of functional explanations, and (iii) it leaves out of sight much of what is achieved by a functional explanation.

I should add a fourth reason: on Canfield’s account functional explanations are explanatory in the sense that they show us how an item is useful to the organisms that have it. However, Canfield fails to explain why showing what an item is good for is explanatory. Hempel, Nagel and Pap employed an inferential theory of explanation. On this theory explanations that fit the covering law model are explanatory because they show us that the phenomenon to be explained was to be expected, in the light of the explaining facts. Canfield argues that functional explanations do not fit the covering law model. But how are we to account for their explanatory force?

6.6 Horan’s account of the explanatory role of appeals to survival value

6.6.1 Horan’s account of functional explanation

Horan (1989) aims to defend the adaptationist program in sociobiology against the criticism (of Gould 1980) that the explanations provided by adaptationists are nothing but a collection of untestable speculations about how a certain trait might benefit its possessors. According to Horan adaptationist explanations should be viewed as functional explanations. They explain “the presence of a pattern of social behaviour in the repertoire of the individuals of a given species” by “the claim that the behaviour pattern has a function”, that is the claim that behaviour “enhances the fitness of an individual who engages in it” (p. 135). According to Horan, sociobiologists have used two methods to provide evidence for such function attributions, namely optimality analyses and the comparative method. Horan argues that adequate evidence for function attributions can be supplied by the comparative method, but not by optimality analyses. As I announced in section 6.3.1 the methods to provide evidence for claims about survival value deserve a special study, which I will not undertake at this place. I restrict myself to Horan’s account of functional explanation.

Horan emphasizes that functional explanation should not be confused with evolutionary explanations. Functional explanations appeal to the current survival value of a trait and explain why that trait is maintained in the species. Thus, they are “forward looking”. Evolutionary explanations are historical or “backward-looking”. They explain the origin of a trait in terms of its past career, including the mechanisms that produced the change. Horan’s account of functional explanation in sociobiology is derived from Cohen’s (1978) account of functional explanation in Marx's theory of history. According to both these accounts functional explanations are a species of so-called "consequence explanations". An example is the following explanation of the monogamous behaviour of many species of small birds:

> (1) If monogamy has a function for individuals in certain species of small birds in environments in which food is scarce, then individuals of these species will be monogamous
>
> (2) Monogamy has a function for these individuals in this kind of environment
>
> --------------------------
>
> (3) Therefore, in this kind of environment, individuals of these species will be monogamous (Horan 1989: 136/7)

This example is made up by Horan herself. She does not substantiate her account with real examples. Neither does she give bibliographic references.

According to both Cohen and Horan consequence explanations are explanatory because they conform to the covering law model. They differ from causal explanations in two ways. First the covering law (1) is a "consequence law" instead of a causal law. Second the initial conditions state at least one "functional fact" (2). Consequence laws tell us that a certain cause is "brought about" or "induced" (p. 136) by the fact that that cause has a certain effect. In sociobiology they are expressed by a functional statement of the following form:

> If trait *T* has a function for individuals of species *S* in environment *E*, then individuals of species *S* will possess trait *T* in *E* (Horan 1989: 136).

According to Horan such laws are supported by evolutionary theory. (I return to that issue in the next section).

A functional fact is an event which is nomically sufficient for the occurrence of its cause (and, hence, which occurs simultaneous with or precedent to the cause). Functional facts are expressed by functional statements of the form:

> Trait *T* has a function for individuals of species *S* in environment *E* (Horan 1989: 136).

Statements of this form say that a certain trait has survival value. Horan argues that it is not possible to show *that* a trait has survival value without showing *how* it does so. In her own words:

> Until one has a well-confirmed hypothesis about *how* a trait enhances fitness, i.e. a hypothesis about its proximate function, there is no reason whatsoever to believe *that it does* enhance fitness, i.e., that it has an ultimate function (Horan 1989: 140).[17]

---

[17] Horan's notion of 'proximate function' confuses attributions of causal roles with attributions of survival value. If a trait has a function the ultimate function of that trait is to enhance fitness. An example of statement

Evidence for such attributions of survival value is to be provided by means of the comparative method.

Hence, on Horan's account the explanatory role of what she calls 'statements of proximate functions' is that they provide evidence for a functional fact stated in a functional explanation. In the next section I discuss the merits of this account.

6.6.2 Evaluation of Horan's account

In this section I argue that Horan's account is unsatisfactory for two reasons. First, Horan's account of consequence laws is confused. Second, Horan does not account adequately of the explanatory role of what she calls "statements of proximate function" (function$_3$ attributions).

Functional explanation vs. evolutionary explanation

Horan repeatedly emphasizes (e.g. p. 135, p. 207-210) that she seeks to do justice to the distinction between functional explanations and evolutionary explanations as this distinction is made by many ethologists following Tinbergen (1963). She seeks to make this distinction in terms of a distinction between "history" and "maintenance". According to Horan both functional and evolutionary explanations are concerned with the process of evolution but they view this process from a different point of view (p. 212). Evolutionary explanations are "historical" or "backward-looking". They explain a trait's "origin" "in terms of the course and dynamics of its past evolution, including possibly, its adaptive significance in past environments" (p. 135). Functional explanations on the other hand are "forward-looking". They explain "why a trait *remains* in the population" (p. 211, emphasis by Horan).

The last two decades the idea that there are two kinds of evolutionary biology, one concerned with history or origin, the other with maintenance or phenotypic existence, gains popularity among evolutionary biologists (see for example Antonovics 1987, Brooks & McLennan 1991, Harvey & Pagel 1991, Reeve & Sherman 1993). I have already discussed Reeve and Sherman's (1993) use of this distinction in section 6.4.4 above. However, it should be noted that the distinction between origin and maintenance can be made in several ways. The evolu-

---

of ultimate function is "The function of monogamy is to enhance fitness" (p. 139). The proximate function of that trait is that what that trait does that enhances its fitness. An example of a statement of proximate function is "The function of monogamy is to ensure adequate provisioning of offspring" (p. 139). As I have discussed in section 6.2 'enhancing the fitness' (in my words: 'having survival value') is essentially comparative. Her example of a proximate function statement, however, is not comparative. Instead, it ascribes a causal role to monogamy. An attribution of survival value would state, for example, that in such and such conditions monogamy has survival value (as compared to polygamy) because in these conditions monogamy ensures the provisioning of offspring better than polygamy.

tionary biologists mentioned above draw the cut at the point at which the relative frequency of the trait concerned becomes stable. Quite often this means at the point where the trait became prevalent. If the distinction is made in this way evolutionary explanations explain how a trait acquired its current relative frequency in the population and explanations of phenotypic existence explain how a certain relative frequency of a trait in the population is maintained after the trait reached that relative frequency. The cut between the origin and the maintenance of a certain trait might also be drawn at the point at which the first variant having that trait showed up in the population. If the distinction is made in that way the manner in which that trait spread through the population after its emergence belongs to its maintenance rather than to its history. Horan does not explicitly state how she makes the distinction between origin and maintenance.

Consequence laws?

Horan's account of consequence laws is confused for three main reasons. First: the consequence law worked out by Horan concerns the evolution of a trait rather than its maintenance. Thus Horan falls prey to the very same sin she criticizes in Wright (1976) and many others, namely the sin of confusing functional explanations with historical explanations. Second, the consequence law worked out by Horan is unacceptable teleological. Third, Horan's consequence law is not supported by evolutionary theory.

Unfortunately, Horan does not incorporate her idea that functional explanations concern the maintenance of a trait explicitly into her account of the structure of functional explanations. She does not use words like "maintained" or "remains" in her account. Instead she uses the phrase "will be". On the covering law model an explanation is an argument, the conclusion of which is a description of the phenomenon to be explained. The conclusion of Horan's example of a functional explanation states that in environments where food is scarce individuals of certain species of small birds will be monogamous. The phrase "individuals of these species *will be* monogamous" (p. 137) sounds more like a prediction than as a description of a phenomenon to be explained. It is true that on the covering law model an explanation is more or less the same as a prediction of the phenomenon to be explained, but this does not mean that the conclusion should be *phrased* as a prediction. Moreover, the conclusion of Horan's example does not say explicitly that the phenomenon to be explained is the maintenance of a certain state of a population rather than the origin of that state. This is asking for confusion.

Indeed, Horan herself falls prey to that confusion. As I quoted in the previous section, according to Horan, the lawlike premise in a functional explanation has the following form:

> If trait *T* has [survival value][18] for individuals of species *S* in environment *E*, then individuals of species *S* will possess trait *T* in *E* (Horan 1989: 136).

for example

> If monogamy has [survival value][19] for individuals in certain species of small birds in environments in which food is scarce, then individuals of these species will be monogamous (Horan 1989: 136/7)

According to Horan such "consequence laws" are supported by evolutionary theory:

> Where the causes of interest are patterns of social behavior, and the effect of importance is enhanced fitness, the elaboration of the mechanisms by which causes are induced by their effects is supplied by evolutionary theory. If, in a given environment, a certain behavior would increase individual fitness, then the chances of reproductive success are greater from individuals who engage in that behavior than for individuals who do not. As a result, more copies of the genetic factors responsible for the disposition to display that behavior will come to be present in the gene pool of the next generation. Hence, if a behavior pattern would increase individual fitness, individuals will come to display that behavior (Horan 1989: 136).

The last sentence of this quote makes clear that according to Horan evolutionary theory supports a consequence law which states that

> if a behaviour pattern would increase individual fitness then individuals will come to display that behaviour.

I take it that the alleged law about monogamy is an instance of this more general "law". It is not clear why her example appeals to this special "law" rather than to the more general one. Anyway, Horan's account of consequence laws is unsatisfactory, for several reasons. A minor one is the lack of a quantifier before "individuals". More serious is the lack of time indicators. The phrase "more copies" in the sentence before the last one in the quote suggests that the quantifier before "individuals" should be 'more'. If an appropriate time indicator is added one gets the following "law":

> if a behaviour pattern would increase individual fitness then in the course of time more and more individuals will come to display that behaviour

more precisely:

> if a behaviour pattern would increase individual fitness then the number of individuals that display that behaviour will increase from generation to generation.

---

[18] Horan uses the words "has a function".

[19] Horan uses the words "has a function".

But perhaps Horan prefers another quantifier and the alleged law should be read as:

> if a behaviour pattern would increase individual fitness then given enough time many/most/all individuals will come to display that behaviour.

In any case, the "law" is phrased as a statement which predicts that under certain conditions a certain state "will come" about. This sounds more like a statement about origin than about maintenance. The use of consequence laws in functional explanations, therefore, violates Horan's distinction between functional and evolutionary explanations. Moreover, the condition under which the state to be explained emerges is the condition that a certain trait *would* have a beneficial effect on its bearers. This is an unacceptable form of teleology.

Horan adds that in general consequence laws are false. It is quite easy to imagine traits that would benefit their bearers but which are absent in those potential beneficiaries. For example, it would benefit fishes to have a third eye, but no fish has. According to her, this is because consequence laws apply only to traits that are "available" (p. 137) to the individuals concerned. As is shown by the fact that no fish has a third eye, third eyes are not available to fishes. However, as is shown by the fact that some small birds are monogamous, monogamy is available to small birds and the corresponding consequence law is true.[20] Horan does not explicitly modify her account of consequence laws but at this point it seems that she holds that the law supported by evolutionary theory is this:

> If a behaviour pattern would increase the fitness of individuals of a certain species *s* and if that pattern is available to individuals of species *s* than in the course of the time (?) more/many/most/all (?) individuals of species *s* will come to display that behaviour.

Horan's addition reinforces the impression that consequence laws are concerned with origins. After all, if the "law" were concerned with the maintenance of existing traits there would be no need to add the requirement that the trait is available. Existing traits are obviously a subset of the available traits. Note also that the modified law is unacceptable teleological in character, just as the original one. Anyway, this law is not supported by Darwinian evolutionary theory. On the contrary, Darwin's theory put an end to the Natural Theologist's idea that the effects a trait *would* have explain why that trait was brought about. According to evolutionary theory the effects important in the evolution of a certain trait are the effects that trait *had* on the relative fitness of their bearers in the past. So much for Horan's account of consequence laws.

---

[20] Actually, Horan says "if it is the case that the fitness of monogamous individuals is greater than that of their polygamous or polygynous conspecifics, the consequence law would be true" (p. 137). This is even more confused than my restatement of here account: for a conditional to be true it is not needed that the antecedent is true.

The explanatory role of statements of proximate function

Horan distinguishes two kinds of function attributions: statements of ultimate function, for example “the function of monogamy is to enhance fitness” and statements of proximate function, for example “the function of monogamy is to ensure adequate provisioning of offspring”. A statement of ultimate function states that a certain trait has a function (that it enhances fitness, or that it has adaptive significance). A statement of proximate function states what that function is (what that trait does to enhance fitness, what its adaptive significance is). In other words it specifies “the causal role that trait plays in an individual’s survival and reproductive success” (p. 141). On Horan’s account functional explanations combine a consequence law with a statement of a functional fact to derive the conclusion that individuals of a certain species will have a certain trait. A statement of a functional fact is a statement of ultimate function. Statements of proximate function are not part of the explanation. Their explanatory role is to support the statement of ultimate function which does the explanatory work.

This account complete ignores both the structure of explanations that appeal to survival value as they are given by biologists and the way in which such explanations work. As my examples in section 6.5.2 show such explanations do not combine a statement simply saying that the trait to be explained has survival value with a consequence law (what ever that may be). Moreover they do not work by producing the expectation that a certain trait will be present, they show in detail why a certain trait is useful (why it has survival value). It is not entirely clear what Horan means by a statement of proximate function. From her example, one might guess that a statement of proximate function states (1) that the ability to perform a certain task (e.g. the capacity to ensure adequate provisioning of the offspring) is influenced by the manner (e.g. monogamous / polygamous) in which a certain type of behaviour (e.g. staying with mates) is performed, and / or (2) the task (provisioning the young) is better performed if the behaviour (staying with mates) concerned has the character it has (monogamous) than if it has some other character (polygamous). The examples I gave in (among other places) section 6.5.2 show that statements of this kind are an integral part of the explanation rather than support for one of the statements that constitute the explanation. The explanatory role of such statements is to help to generate the insight provided by the explanation, rather than to convince the audience that one of the statements that constitute the explanation is true.

6.6.3 Conclusion

Let us now see whether Horan’s idea that functional explanations fit the covering law model, can be saved by replacing the consequence law in her account by another law that appeals to survival value. One candidate is, of course, the following principle of natural selection:

> if some (but not all) individuals of a certain species *s* have a certain trait, *T*, and if the fitness of individuals of species *s* having *T* is higher than the fitness of individuals of species *s* lacking *T* then the relative frequency of individuals of species *s* that have *T* will increase.

It is assumed that the differences are inheritable. The problem is that this statement is not a law in the sense of the covering law model. For it is not universally true: genetic linkage and genetic drift might prevent the trait that confers the greater fitness to its bearer from increasing its share in the population. The point is that whereas the past effects of a certain trait on the relative fitness of individuals with that trait do explain the increase of the relative frequency of that trait in the population, there is no *law* which states that increase in relative fitness leads to increase in relative frequency. Hence, the covering law model fails to make sense of such explanations.

In Horan’s elaboration, the lawlike premise predicts that under certain conditions the number of individuals having a certain trait will become prevalent, or, at least, increase their share in the population. Horan’s remark that functional explanations are concerned with the maintenance of a trait suggests another elaboration of the first premise, namely as a statement about the maintenance of a trait after it has become prevalent:

> If trait *T* is prevalent among individuals of species *s* in *E* and if trait *T* confers higher fitness to individuals of species *s* in *E* than all available alternatives to *T* would do, then trait *T* will remain prevalent in future generations of species *s* in *E*.

This elaboration differs from Horan’s elaboration in that it is restricted to existing (more specifically: prevalent) traits rather than to available ones. The consequence of this principle states that the trait concerned will remain prevalent rather than that “individuals will come to have the trait concerned”. Furthermore, it is explicitly required that trait *T* confers a greater fitness to individuals of species *s* in *E* than all available alternatives would do. This principle is not a law in the sense of the covering law model for it is not universally true: the fittest trait might loose its prevalence due to genetic drift.

Nevertheless, the idea that appeals to survival value might provide insight in the way in which a certain trait is maintained in the population is most promising. There are many examples of explanations which attempt to do this. I shall call such explanations ‘equilibrium selection explanations’. Whereas evolutionary selection explanations explain how a certain change in a population was brought about, equilibrium selection explanations explain why the population remains in a certain state. Sober (1984) gives an important account of such explanations.[21] Reeve & Sherman (1993) suggest that such an account accounts for the explanatory force of appeals to adaptedness$_4$ (as I have discussed in section 6.4.4). Sober’s main example is

---

[21] Sober does not use the term ‘survival value’. He talks of explanations that appeal to fitness or adaptedness (that is adaptedness$_4$).

Fisher's (1930) explanation of why, the ratio of males to females in most species is 1:1 (at least at the zygote stage). This ratio was a riddle to Darwin who observed that a parent who produces ten daughters has the same number of offspring as a parent who produces five sons and five daughters (Darwin 1886: 259). Fisher points out that the riddle can be solved by taking the third generation into account. Suppose that the second generation consists of f females and m males which together produce N offspring. This means that the number of offspring produced by a female of the second generation equals to N/f and that of a male of the second generation to N/m. As a result, if the sex ratio differs from 1:1, an individual of the first generation will have more grandoffspring if it produces more offspring of the minority sex. This means that if the sex ratio differs from 1:1 a variant that produces more offspring of the minority sex will rapidly spread through the population, as a result of which the minority sex increases its share in that population. This continues until the sex ratio is equal to 1:1. At this point there is no advantage in producing more of one sex. The 1:1 sex ratio is therefore a state that will be maintained by natural selection.

Although the term 'survival value' is not mentioned, Fisher's explanation clearly appeals to the survival value of producing more offspring of the minority sex (as compared to producing the same number of offspring of both sexes). It is certainly the case that biologists appeal to survival value to explain why a certain trait remains prevalent in the population. This use of the notion of survival value accounts in part of the explanatory use of attributions of survival value.

However, in design explanations appeals to survival value have another use. In design explanations the fitness of the real organism is compared to that of a hypothetical organism. The hypothetical organism might be a variant that can easily turn up in the population, but in many cases the comparison is between a real organism and a hypothetical organism that cannot easily arise from the current population. For example, Krogh compares vertebrates with circulatory systems with similar hypothetical organisms that have to rely on diffusion alone. Similarly, Tinbergen compares eggs with natural colours with eggs that are painted white. Habibi c.s. compare sand gazelles and mountain gazelles with hypothetical gazelles that behave like gazelles of one species and live in the conditions of the other species. Such comparisons yield information about the way in which the individual hangs together, apart from possible information about the way in which the traits concerned are maintained in the population. The accounts of Sober and Reeve & Sherman do not account for this insight in relations at the individual level.

6.7 Conclusion

In this chapter I have analyzed the distinction between function as causal role (function$_2$) and function as survival value (function$_3$). Attributions of a causal role are concerned with the contribution of a certain item or behaviour to the ability to perform a certain task. The causal role of an item is its position in a system that performs that task. This position does not depend on the environment, neither on other possible ways to perform that task. Claims about survival value are concerned with the effects on fitness of the manner in which a certain task is performed. As such they depend on a preceding attribution of a causal role. Claims about survival value compare the way in which a certain task is performed in real organisms with other possible ways in which that task might be performed. Whether one way to perform a certain task is better than another, is, of, course, dependent on the environment. This means that the survival value of a certain trait is relative to the environment. The ultimate criterion for what counts as better is the fitness of the organism that have it.

Explanations that appeal to survival value rely heavily on counterfactual comparison. Philosophers have made the following objections against an account of function in terms of counterfactual comparison: (1) there is no unique way to determine a counterfactual situation, (2) counterfactual comparison assumes that the laws of nature do not apply, (3) counterfactual comparison is complicated and speculative. Against the first objection I have argued that a unique reference situation is required only if one wants to use counterfactual comparison too single out causal roles, but not if one uses it to asses survival value. After all, claims about survival value are relative to a reference situation. Against the second objection I have argued that one can make counterfactual comparisons without assuming that the laws of nature do not apply. On the contrary the laws of nature allow us to make reliable comparisons. Against the third objection I have argued that the application of physics and chemistry allows for reliable counterfactual comparison.

The main challenge to a philosophical account of explanations that appeal to survival value, is to explain how comparison with hypothetical organisms that have never existed can provide insights in real organisms. The present accounts within the survival value approach fail to do this.