Explanation Without A Cause

Verklaren Zonder Oorzaken te Geven

A.G. Wouters

Z∃NO

Zeno, The Leiden-Utrecht Research Institute of Philosophy
Heidelberglaan 8
3584 CS Utrecht
The Netherlands

Quastiones Infinitae

PUBLICATIONS OF THE ZENO INSTITUTE OF PHILOSOPHY

VOLUME XXIX

ISBN 90-393-2228-7 (ISSN 0927-33954)

Cover design: Het Bona-team

Explanation Without A Cause

Verklaren Zonder Oorzaken Te Geven

(met een samenvatting in het Nederlands)

Proefschrift ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. dr. H.O. Voorma,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen
op woensdag 3 november 1999 des middags te 12.45 uur

door

Arno Gerhard Wouters

geboren op 5 augustus 1956 te Amsterdam.

promotoren: Prof. dr. A.A. Derksen, Faculteit der Wijsbegeerte, Katholieke Universiteit Nijmegen
Prof. dr. J.A. Bergstra, Faculteit der Wijsbegeerte, Universiteit Utrecht

The investigations were supported by the Foundation for Philosophical research (SWON), which is subsidized by the Netherlands Organization for Scientific Research (NWO).

Table of Contents

3 Biologists about function 57

4 The classical attempts 69

5 The causal role theory 95

Chapter 1: Introduction

Philosophers of science with an interest in biology have always been intrigued by the fact that biologists often explain phenomena by appealing to functions rather than to causes. Biologists are perceived as people who explain why certain organisms have certain parts or organs or why they perform certain activities by citing the functions of those items or activities. It is, for example, generally thought, among philosophers of science, that biologists explain the presence of hearts in many organisms by appeal to the heart's function to propagate the blood. Such explanations are called 'functional explanations'. The question how it can be explanatory to cite functions has become the classical philosophical issue in regard to biological explanation.

A related issue is the meaning of the notion of 'biological function'. Analyses of this notion have served quite different philosophical agendas. In the last decade philosophical analyses of function have become increasingly important in the philosophy of language and the philosophy of mind. The main impulse for this outburst was Ruth Millikan's *Language, Thought and Other Biological Categories* (1984). In this highly original and important work Millikan employs a certain theory of function (the so-called 'etiological theory') to solve many problems in the philosophy of mind and language. My agenda is different. I seek to understand the actual practice of explanation in functional biology.

Functional biology is that part of biology which is concerned with the structure, activity and behaviour of individual organisms. Functional biologists investigate how the different organisms that populate our planet are built, how they behave and how they work. Of course, functional biologists are not only interested in *describing* the parts, processes and behaviours of the organisms they study, they also seek to *explain* the way in which those organisms are built, the way in which those organisms behave, and the way in which those organisms work. It is those explanations in which I am interested.

The philosophical study of explanation is often viewed as an exercise in conceptual or linguistic analysis. On this view, the philosopher's main task is to provide a definition of the notion 'explanation' that applies to all or most of the examples that intuitively count as examples of good explanations. A theory of explanation should be criticized by providing counterexamples showing that certain explanations that are intuitively good explanations are not explanatory according to the criticized theory or that the theory does not exclude certain pieces of reasoning which are intuitively not explanatory. In regard to functional explanations the main question is whether and how appeals to function are explanatory. This question is to be answered by discussing how such appeals relate to one's favourite definition of explanation.

I am interested in a more substantial issue than the one above, namely the question what is achieved by a functional explanation (what do the pieces of reasoning which biologists call ‘functional explanation’ add to our knowledge?). I take the question ‘how can it be explanatory to appeal to a function?’ as meaning ‘what do biologists learn from accounts that appeal to functions on top of the descriptions employed in those accounts?’, ‘why is it useful to attempt an account in functional terms?’. That is, I am not interested in the question whether or not so-called ‘functional explanations’ are really explanations but in the question what such pieces of reasoning add to our knowledge.

In this introductory chapter I shall discuss how the problem of functional explanation entered contemporary philosophy of science in the context of the deductive nomological model of scientific explanation and what the current approaches to this problem are (section 1), summarize the main issues often discussed in relation to functions and functional explanations (section 2), and outline my own views on this subject (section 3).

1.1 The problem of functional explanation

The problem of function has entered contemporary philosophy of science in the context of the deductive-nomological model of scientific explanation (Hempel & Oppenheim 1948). According to this well-known model explanations deduce a statement describing the phenomenon to be explained from a combination of statements expressing general laws and statements describing initial conditions. This model was thought to capture the principal characteristics of explanation in the physical sciences. The deductive-nomological style of explanation was also found in biology. An example is the explanation of a certain regularity in the progeny of hybrid peas obtained by crossing inbred peas with round seeds with inbred peas with wrinkled seeds (the ratio of plants with round peas to plants with wrinkled peas in this so-called F_2 generation is approximately 3 : 1). This regularity is explained by deducing it from certain general principles of the Mendelian theory of heredity together with assumptions about the genetic constitution of the parent peas (Nagel 1961: 18). However, the proponents of the deductive-nomological model also noted that biologists frequently employ other “types” or “methods” of explanation that appear to be very different from the types or methods of explanation used in the physical sciences, at least at first sight. One such type or method was called “functional explanation” (Nagel 1961), “teleological explanation” (Nagel 1961, Pap 1962, Canfield 1964), or “functional analysis” (Hempel 1959). (I will use the term ‘functional explanation’.) The explanation of the heart or the beating of the heart by appeal to its function to circulate the blood has become the standard example of such a functional explanation.

According to the proponents of the deductive-nomological model, functional explanations purport to explain items or activities in terms of the functions that item or activity has for the organisms which have that item or perform that activity. They are often characterized by the occurrence of teleological expressions such as ‘the function of’, ‘the role of’, ‘serves as’, ‘in order to', 'for the sake of', 'for the purpose of'. Functional explanations as they are used in biology have neither anthropomorphic nor vitalistic implications. That is, when biologists state that a certain item or activity has a function they do not imply that those items or activities result from design by an intellectual being (anthropomorphism) or from the activity of extra-physical, vital forces (vitalism). Yet, functional explanations differ considerably from explanations in the physical sciences. As Nagel put it in his *Structure of Science*:

> It would surely be an oddity on the part of a modern physicist were he to declare, for example, that atoms have outer shells of electrons in order to make chemical unions between themselves and other atoms possible (Nagel 1961: 401)

The absence of functional explanation in the physical sciences and its presence in biological sciences posed a two-edged challenge to the proponents of the deductive-nomological model in post-war philosophy of science. On the one hand they had to face the apparent consequence of their theory, namely that teleological language in biology was a sign of its immaturity. In the physical sciences teleology has been banned since the seventeenth century. Because of the prestige of physical sciences and because of the tendency of philosophers to regard physics as the paradigm of science many philosophers as well as scientists tended to view explanations that use teleological expressions as obscurant and suspicious. However, philosophers of science like Carl Hempel and Ernest Nagel, were very much aware of the success of the functional mode of explanation in biology and in the social sciences. This made it difficult for them to dismiss functional explanation simply as immature.

On the other hand they had to face those who argued that the differences between the mode of explanation in the physical sciences and the mode of explanation in the life sciences reflected a difference in the nature of the phenomena studied. The modes of explanation appropriate in physics, it was urged, are not appropriate in biology. This idea conflicted with the idea of the unity of science embraced by the philosophers of science of that time.

As a result philosophers of science with an interest in biology saw a two-fold task for themselves. Their first task is to show that the use of teleological language in biology is innocuous because teleological statements can be translated without any loss of content into statements that do not contain teleological expressions. Their second task is to account for the explanatory use of teleological expressions in terms of the deductive-nomological model.

The main obstacle to this program (at least as it was initially perceived) is the so-called 'problem of functional equivalents', that is the existence of different ways to perform a certain function (think of alternate devices such as artificial hearts that might circulate the blood). The classical analyses in this context are those of Hempel (1959) and Nagel (1961, 1977). Both Hempel and Nagel employed an inferential conception of explanation. On this view of explanation, explanations work by showing that the phenomenon to be explained was to be expected in virtue of the explaining facts. Hempel and Nagel differed on the issue whether or not function attributions allow one to infer the presence of the item to which the function is attributed. According to Nagel, given a certain form of organization, a certain item is necessary to perform a certain function (that is there are no real functional equivalents). Therefore, given the function (and the form of organization) one may infer the presence of the functional item and, consequently, functional explanations are really explanatory. According to Hempel on the other hand the existence of functional equivalents shows that functional traits are not really necessary (there are real functional equivalents). An organism must perform certain functions, but since there are different ways to perform those functions, the functional item itself is not necessary. Attributions of functions are explanatory only in the limited sense that they allow one to infer the presence of one of the several items of an ill-defined class of items capable of performing a certain function. The main scientific value of the pieces of reasoning which are often called ‘functional explanation’ is heuristic: they lead to the discovery of new phenomena.

Most recent accounts of function and functional explanation have abandoned the idea that functional explanations explain by deducing a sentence stating that organisms of a certain taxon (must) possess a certain item or perform a certain activity, from a combination of sentences stating initial conditions and sentences stating general laws. It was probably John Canfield (1964) who first distinguished functional explanations from explanations that fit the covering law model. According to Canfield, functional explanations do not subsume the presence of an item under a general law; instead they specify what that item does that is useful to the organisms that have it.

> The latter [explanations fitting 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][1] in biology [...] 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).

It is important to distinguish two issues in the rejection of the deductive nomological model. One concerns the structure of functional explanation, the other their nature.

In regard to the first issue many recent writers appear to agree with Canfield that functional explanations consists of a single function attribution in answer to a question of the form ‘why do organisms of taxon *t* have item/trait *i*?’. For example, Millikan states:

> if you want to know why current species members have T the answer is, very simply, because T has the function F (Millikan 1989a: 174).

---

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

On the issue of the nature of functional explanations both Pap (1962) and Nagel (1977) are of the opinion that functional explanations are not of a causal nature. Pap emphasizes that

> although [functional explanations][2] have the same logical form as explanations in terms of efficient causes, it would be highly misleading to call them “causal.” (Pap 1962: 361)

Nagel discusses an example of a functional explanation which has roughly the following form:

> (1) This plant performs photosynthesis.
>
> (2) Chlorophyll is necessary for plants to perform photosynthesis
>
> ------------------
>
> (3) Hence, this plant contains chlorophyll.

This explanation conforms to the deductive-nomological model. The second premise is lawlike, the first one states initial conditions. However,

> the performance of photosynthesis is not an *antecedent* condition for the occurrence of the chlorophyll, and so the premise is not a causal law. Accordingly, if the example is representative of [functional explanations][3], such explanations are *not* causal—they do not account causally for the presence of the item to which a function is ascribed (Nagel 1977: 300)

He adds to this that

> inquiries into effects or consequences are as legitimate as inquiries into causes or antecedent conditions (Nagel 1977: 301)

On the inferential conception of explanation, endorsed by the proponents of the deductive-nomological model, explanations need not be causal to be explanatory, provided that they make it plausible that one should expect the presence of the phenomenon to be explained. The causal conception rejects this view of explanation in favour of the idea that an explanation should show us how the phenomenon to be explained is brought about. Robert Cummins (1975) applies this idea to functional explanations and draws the conclusion that it is impossible to explain the presence of an item by appeal to its function:

> To attempt to explain the heart’s presence in vertebrates by appealing to its function in vertebrates is to attempt to explain the occurrence of hearts in vertebrates by appealing to factors that are causally irrelevant to its presence in vertebrates. Even if it were possible, as Nagel claimed, to *deduce* the presence of chlorophyll from the occurrence of photosynthesis, this would fail to *explain* the presence of chlorophyll in green plants in just the way deducing the presence and height of a building from the existence of its shadow would fail to explain why the building is there and has the height it does (Cummins 1975: 745/6)

---

[2] Pap uses the term 'teleological explanations'.

[3] Nagel, at this point, uses the expression 'explanations of function ascriptions'

Cummins, for that reason, rejects the assumption that “the point of functional characterization is to explain the presence of the item (organ, mechanism, process or whatever) that is functionally characterized” (p. 741). According to Cummins what we can explain and what we do explain by appeal to functions is the activity or the capacity of a system of which the item to which the function is attributed is a part. In his view, functional explanations explain how a system is able to perform a complex task by pointing to the capacity of the parts of that system to perform a series of subtasks that add up to the original capacity. For example, a functional explanation of an organism’s capacity to circulate oxygen would explain this capacity as the result of the combination of the capacity of the blood to carry the oxygen, the heart to pump the blood around, and the blood vessels to direct the blood from the lungs or gills (where the blood is aerated) to the organs (where the oxygen is consumed) and back. The attribution of the function of propagating the blood to the heart serves to explain the capacity to circulate the blood but it does not serve to explain the presence of the heart.

The distinction between the inferential and the causal conception of explanation has been driven home by Wesley Salmon in his “Why Ask, ‘Why?’?” (1978). Salmon’s *Scientific Explanation and the Causal Structure of the World* (1984) is the classical statement of a causal theory of explanation. As Salmon (1984: 15-20, 1989: 119-121) puts it the inferential conception is oriented towards the notion of nomic expectability. On this view an explanation is an argument that reveals that the phenomenon to be explained was to be expected in virtue of the laws of nature and the conditions in which that phenomenon occurred. The causal conception, on the other hand, focuses on how phenomena fit into the causal structure of the world. On this view, explanations reveal the mechanisms that produce the phenomenon (transition, state or property) to be explained.

In the causal view of explanation the problem how it can be explanatory to cite the function of an item takes a form different from the one it had on the inferential theory. Whereas on the inferential conception the main problem of functional explanation is the problem of functional equivalents, on the causal conception the main problem of functional explanation is the problem that function attributions apparently refer to effects rather than to causes. This problem has been formulated most clearly by Karen Neander

> The general prima facie problem with [functional explanations][4] is often said to be that they are ‘forward-looking’. [Functional explanations] explain the means by the ends [...], and so the explanans refers to

---

[4] Where I have put "functional explanations" the author uses the expression "teleological explanations", this category includes two kinds: intentional explanations and functional explanations. She adds that the solution of the problem is transparent in cases where an intentional agent is involved (that is in the cases of intentional explanations and of functional explanations that appeal to the functions of artefacts). The more intransigent problem arises where there is no intentional agent involved, as with biological function.

> something that is an effect of the explanandum, something that is forward in time relative to the thing explained. [...] Indeed, because [functional explanations] seem to refer to effects rather than prior causes, it looks at first sight as though backward causation is invoked. [...] The prima facie problem gets worse, if that is possible, because many [...] functional effects are never realized (Neander 1991b: 455/6).

Sandra Mitchell makes a similar point:

> The classic[5] philosophical puzzle regarding functions is how could appeal to a consequence of the presence of a trait explain why the trait is in fact present, since the functional consequence is temporally later and so cannot be causal in a straightforward sense (Mitchell 1993: 249)

As I mentioned above, Cummins tries to solve this problem by maintaining that functional explanations are not intended to explain the presence of the item to which the function is attributed. According to him functional explanations explain a capacity of a system of which the item to which a function is attributed is a part. Cummins emphasizes the difference between explanations that explain changes in the state of a certain system (transition explanations) and explanations that explain the properties of a certain system (property explanations). According to him the covering law model applies to transition explanations but not to property explanations. Functional explanations are a special kind of property explanations, in which the property to be explained is a complex capacity and the explanation proceeds by analyzing that complex capacity into a number of less complex capacities. Cummins tends to restrict the term ‘causal explanations’ to explanations that explain transitions. However, functional explanations in Cummins’s sense are causal explanations in the sense in which Salmon uses the term ‘causal explanation’: they reveal the processes and mechanisms that bring about a certain phenomenon (in the case of functional explanations the phenomenon to be explained is a capacity). I shall use the term ‘causal explanation’ in Salmon’s sense. This means that, on my account, functional explanations in Cummins’s sense are of a causal nature.

Whereas Cummins solves the problem of functional explanation on the causal account by denying that function attributions are used to explain the presence of the item to which the function is attributed, most other authors have attempted to reconcile the idea that functional explanations explain the presence of the item to which the function is attributed with the causal account. One of the first attempts to formulate such a theory of function is the theory of Larry Wright (1973, 1976). According to Wright:

---

[5] Apparently, Mitchell is herself not aware that this wording of the problem of functional explanation is peculiar to the causal conception of explanation and she assumes, mistakenly, that Hempel resolved this problem by arguing that functional explanations do not meet the conditions of legitimate explanations

> The treatments we have so far considered have overlooked, ignored, or at any rate failed to make, one important observation: that [function attributions][6] are—intrinsically, if you will—explanatory. Merely saying of something, *X*, that it has a certain function, is to offer an important kind of explanation of *X* (Wright 1973: 154),

He adds to this that the explanatory force of function attributions is not merely that they state “what the functional thing is good for” (as Canfield has argued). Function attributions are “explanatory in a rather strong sense” (p. 155): they explain “how the thing with the function got there” (p. 156). Hence, they are explanatory on a causal view of explanation.

Current theories of functions and functional explanations come in roughly three kinds.[7] According to the *etiological theory* (for example Millikan 1984: 17-49, Millikan 1989b, Mitchell 1989, Brandon 1990: 184-189, Neander 1991a,b, Millikan 1993a)[8] function attributions specify the effects for which a certain trait was selected in the past. These effects are the functions of that trait (many proponents of an etiological account speak of “proper functions” where I would use “functions”). For example, hearts are said to have the (proper) function of

---

[6] Wright uses the words "functional ascriptions"

[7] See Amundson and Lauder (1994). Achinstein (1977) and, in his trail, Reznek (1987) distinguish three kinds of explications of (F) "The function of X (in Y) is Z": (1) (the "good-consequences doctrine" or "evaluative theory"): X does Z and X's doing Z contributes to the good of Y (for example Hempel 1959, Canfield 1964, Sorabji 1964, Lehman 1965a, Ruse 1973, Hull 1974, Woodfield 1976); some accounts in this class identify specific goods (e.g. Hempel and Ruse who define the good as "adequate working order" respectively "survival and reproduction") others (e.g. Sorabji and Woodfield) leave this open; (2) (the "goal doctrine" or "teleological theory"): X does Z and X's doing Z contributes to some goal of Y (for example Nagel 1961, Boorse 1976); and (3) (the "explanation doctrine" or "etiological theory"): X does Z and X doing Z makes a causal contribution to X's occurrence in Y (for example Ayala 1970, Wright 1973, Bennett 1976, Levin 1976, Wright 1976). Apart from some differences on who belongs to which approach, their classification differs from mine on three points: (1) Achinstein and Reznek fail to take Cummins's approach into account; (2) Achinstein's and Reznek's descriptions of the etiological account are ambiguous about the level of explanation (does Y refer to an individual or to a population?); (3) I do not distinguish a separate "goal-doctrine". Depending on what kind of goals are involved I treat the different goal-doctrines as variants of either the causal role theory or the survival value approach.

[8] Forerunners of the etiological theory as it is defined here are Ayala (Ayala 1970) and Wright (1973, 1976) according to whom functions are determined partly, but not wholly by a selection history. Wimsatt's (1972) account, also, has certain elements of an etiological theory but its main tenure is that of a survival value account. A related approach is that of Reznek (1987: 117) according to whom the functions of the trait are the effects that make a causal contribution to the continued presence of the trait in the population via the mechanism of natural selection.

propagating the blood, if and only if propagating the blood is what hearts did that caused them to be favoured by past natural selection. The etiological theory of function attributions fits nicely with the causal conception of explanations. On the etiological view, function attributions are intrinsically explanatory in that they specify the factors that account for a trait's presence in a certain population. They specify the effects of a certain trait that were causally effective in the evolution of that trait. Among evolutionary biologists and among naturalistic philosophers of science the etiological view "is fast becoming the consensus" (Neander 1991a: 168).

A second approach identifies functions with effects that currently make a causal contribution to the survival and reproduction of an individual organism (for example Canfield 1964, Wimsatt 1972, Ruse 1973, Bigelow & Pargetter 1987, Horan 1989). On this view the function of the heart is said to propagate the blood because that is how hearts currently contribute to the survival and reproduction of the organisms that have hearts. I shall call this approach the *survival value approach* to functions. The proponents of the survival value approach differ in the way in which they account for the explanatory force of function attributions. Canfield denies that function attributions explain the presence of the item to which the function is attributed, function attributions merely show us how a certain item is useful. Ruse and Horan employ the inferential theory; they attempt to account for functional explanations as deductive arguments that allow one to infer the presence of the item to which the function is attributed. Wimsatt, too, employs the inferential theory, but according to him functional explanations have a statistical rather than a deductive character. Bigelow & Pargetter attempt to account for functional explanations in terms of the causal theory by emphasizing that a function is a disposition to have a certain effect, rather than that effect itself.

On the third approach, function attributions describe the role of a certain part or activity in maintaining a certain capacity of a system to which it belongs. On this account to say that hearts have the function to propagate the blood is to say that propagating the blood is what the heart does that accounts for the organism's capacity to circulate the blood. This approach is best known from Cummins's "functional analysis" (1975). The idea is also present in embryonic form in Nagel's (1961, 1977) work:

> [Functional explanations][9] make evident one role some item plays in a given system (Nagel 1977: 300).

Following Neander (1991a: 181) and Amundson & Lauder (1994), I shall call this approach the *causal role theory* of function. On the causal role theory, function attributions serve to explain a complex capacity of a system of which the item to which the function is attributed is a part. Such an account is explanatory on a causal theory of explanation.

---

[9] At this point Nagel uses the expression "explanations of function ascriptions".

The inferential theory of explanation and the causal theory of explanation offer two different answers to the question what is learned from explanations. On the inferential theory of explanation explanations tell us that the phenomenon to be explained is to be expected in view of the laws of nature and the conditions applying to that phenomenon. On the causal theory of explanation explanations reveal the mechanisms that produce the transition, state or property to be explained. The different philosophical theories of functional explanation explain whether, why and how appeals to function are explanatory (or not) on some view of explanation. Philosophers who think that appeals to function are not explanatory on their favourite theory of explanation tend to think of the scientific value of such appeals as heuristic rather than explanatory; by this they mean that the scientific value of this kind of reasoning is to be located exclusively in the process of the discovery of new facts, rather than in the explanation of the discovered facts.[10]

As I said in the introduction to this chapter, I aim to understand the practice of explanation in functional biology. I seek to answer questions like ‘why is it useful to attribute functions to the parts and behaviours of organisms?’ and ‘what do biologists learn from the pieces of reasoning in which they appeal to functions?’. This means that the problem of functional explanation in which I am interested is different from the problem of functional explanation on the inferential theory and also from the problem of explanation on the causal theory. My starting point is not a philosophical theory of explanation but the practice of reasoning in functional biology. I aim to understand what the different kinds of reasoning that constitute this practice contribute to the process of enquiry. I focus on reasoning concerning functions. I distinguish several kinds of reasoning concerning functions. My main contention is that functional biologists propose and defend (among other things) accounts which do not show how a certain transition, state or property is brought about, nor do they show that a certain phenomenon was to be expected in the light of the facts mentioned in the explanandum. Yet, they add something to our knowledge in addition to the new discoveries to which such reasoning may lead.[11] I shall call this kind of reasoning ‘design explanation’. An example of a design explanation would be an explanation of the hollow character of the heart by showing that this character is needed to enable the heart to pump blood (if the heart were solid or spongy it would not be able to pump blood). Design explanations are overlooked in contemporary philosophy of science. The problem of functional explanation as I see it, is the problem what design explanations add to our knowledge (in addition to the facts they describe).

---

[10] A recent proponent of the view that functional or teleological language is merely heuristic is Schaffner (1993: ch. 8).

[11] See Resnik (1995) for an account of the role of functional language in biological discovery.

1.2 The main issues discussed in connection with the notion of ‘function’

In the previous section I discussed the problem of functional explanation and outlined the main approaches to this problem. In this section I shall list a number of issues often discussed in philosophical analyses of the notion of function (remember that not all analyses aim to account for the problem of functional explanation). Woodfield (1976) lists three issues that tend to show up in philosophical analyses of functions (Reznek (1987: 99/100) and Pranger (1990: 64) copy this list without much commentary):

> (1) Hempel’s problem. Why is it that only some of the item’s activities are functions, and the others accidental?
>
> (2) Nagel’s problem. Why is it that we ascribe functions to the parts of some systems (like organisms) but not to the parts of others (like the solar system)?
>
> (3) The problem of functional explanation. How can it be explanatory of an item to cite one of its effects? (Woodfield 1976: 108)

Already in Woodfield’s time it was clear that Hempel’s problem (the function/accident distinction) consists of a number of different problems/distinctions that can not always be treated equally. The analyses of Wright (1973, 1976), Millikan (1984) , Horan (1989) and Neander (1991a) added new problems to the agenda.

The following list presents a fairly comprehensive overview of the issues that have been raised with respect to biological function in the form of a list of desiderata.

1) A theory of function should not allow one to ascribe functions to parts of “purely physical” systems such as solar systems (Nagel 1961: 406). The challenge is, of course, to define ‘purely physical systems’.
2) A theory of function should distinguish between activities that are functions (such as the beating of the heart) and activities that are useless side-effects of functional organs (such as heart sounds and pulses) (Hempel 1959).
3) A theory of function should not depict the use other organisms make of the items of a certain organism as functions of those items. It is, for example, not a function of a dog’s long hair to harbour fleas (Ruse 1973: 183).
4) A theory of function should distinguish between effects that are functions and effects that are accidentally useful. Although belt buckles occasionally save their wearers’ life by deflecting bullets, it is not a function of belt buckles to deflect bullets (Wright 1973: 147).
5) A theory of function should not depict the use of existing items for new purposes as functions of those items. It is, for example, not the function of the human nose to support eyeglasses (Wright 1973: 148).

6) A theory of function should distinguish currently functional items from vestiges (like vermiform appendices in humans and vestigial eyes in cave dwellers) (Wright 1976: 87, Griffiths 1993).
7) A theory of function should enable us to attribute functions to items that are incapable of performing their function (e.g. malformed hearts) or do not actually perform it (e.g. most sperm cells) (Millikan 1984, Millikan 1989a, Neander 1991a).
8) A theory of function should not confuse functional explanations and evolutionary explanations (Horan 1989, Godfrey-Smith 1994).

1.3 An outline of my argument

I start with a description of the practice of functional biology (chapter 2). After a short introduction I focus on the notion of ‘function’ (section 2.2). Most of the proponents of a certain account of function assume that within biology the word ‘function’ has a unique meaning. The different proposals are usually seen as rival analyses of that unique notion of function. I shall show by means of examples of real biological research that the word ‘function’ is used in biology in a number of different ways. In relation to the problem of explanation it is important to distinguish at least four kinds of function:

(1) function as activity (function$_1$)—what an organism, part, organ or substance does or is capable of doing;
(2) function as causal role (function$_2$)—the role of a part, organ, substance or behaviour of an organism in maintaining a complex activity or capacity;
(3) function as survival value (function$_3$)—the survival value of a certain part, organ, substance, or behaviour; or of a part, organ, substance or behaviour having a certain character;
(4) function as selected effect (function$_4$)—the effects for which a certain trait was selected in the past and that explain its current presence in the population.

For example, in the case of the heart, the beating of the heart is an activity (function$_1$), the performance of which explains the heart’s causal role (function$_2$) in circulation, namely propagating the blood. The survival value (function$_3$) of having a heart is presumably that it is more efficient to have one organ (the heart) as a source of energy for circulation than to have all big blood-vessels beating.[12] It is highly problematic and speculative to specify why the heart was

---

[12] A good explanation by specification of function as survival value (function$_3$) must of course specify how the heart is more efficient than a system of beating blood-vessels. The survival value (function$_3$) of the system of heart and blood-vessels as a whole is that it satisfies the need for circulation of, among others, oxygen.

selected in the past (function$_4$), but we may guess that the aforementioned efficiency has played a role.

In section 2.3 I show that functional biologists try to answer seven different types of questions. These questions concern:

(1) the structure and activity of the organisms they study;
(2) the causal roles of the parts and behaviours of those organisms;
(3) the mechanisms by means of which these causal roles are performed,
(4a) the survival value of performing these causal roles;
(4b) the survival value of having a certain character;
(5) the development of those organisms in the course of the ontogeny;
(6) the evolutionary history of those organisms, their parts, and their behaviour.

The products of enquiry are respectively:

(1) descriptions of the structure and activity of organisms and their parts;
(2) attributions of causal roles (function$_2$ attributions);
(3) physiological explanations;
(4) design explanations;
(5) developmental explanations;
(6) evolutionary explanations.

I also show that attributions of causal roles are the handle by means of which functional biologists get a grip on their subject matter. They are used in at least three different types of explanations:

1) physiological explanations of the capacities of a system of which the item to which the causal role is attributed is a part (capacity explanations). For example, the insight that the heart propagates the blood helps to explain how the organism is able to circulate oxygen.
2) design explanations of certain aspects (shape, structure, activity etc.) of the part or behaviour to which the causal role is attributed. For example, the causal role of the heart in blood circulation helps to explain why the heart is hollow.
3) explanations of the evolution of the part or behaviour to which the causal role is attributed. For example, the causal role of the heart in oxygen circulation helps to explain why the heart evolved from a two chambered to a four chambered state.

The attribution of a causal role is the first step in these explanations but it does not constitute the complete explanation.

In chapter 3 I relate the description of the practice of functional biology which I presented in chapter 2 to some descriptions of that practice by biologists.

In chapter 4–7 I investigate the extent to which the philosophical theories of function and functional explanation are of use to understand the practice of functional biology.

In chapter 4 I am concerned with the attempts of Hempel and Nagel to account for functional explanations in terms of the inferential theory of explanation. Hempel and Nagel do not give real examples of the explanations about which they talk. This makes it difficult to determine which kind of explanation they have in mind when they talk of “functional explanations”. I suggest that they are concerned with design explanations. I argue that they fail to account for the explanatory force of this kind of explanation (that is they fail to explain what design explanations add to our knowledge).

In chapter 5 I am concerned with Cummins’s causal role theory of function. I argue that this theory offers an excellent account of the notion of function as causal role (function$_2$). Moreover, Cummins offers an excellent account of an important kind of physiological explanation, namely capacity explanation, and of the role of attributions of causal roles in that kind of explanation. His theory fails to account for the role of attributions of causal roles in design explanations and evolutionary explanations.

In chapter 6 I discuss several proposals within the survival value approach. I argue that these proposals tend to confuse the notion of function as survival value (function$_3$) with the notion of function as causal role (function$_2$). Nevertheless, they offer important contributions to an analysis of the notion of survival value (function$_3$). However, the current proposals fail to account for the explanatory force of appeals to survival value (they fail to explain what one learns from an appeal to survival value). In this chapter I also elaborate on my own account of survival value which I started in chapter 2.

In chapter 7 I argue that the etiological theory does not apply to most function attributions in both functional and evolutionary biology. As a result it leaves the explanatory force of all these function attributions unexplained (that is, it tells us nothing about why such function attributions are important in science). This means that this approach is irrelevant to my project, which aims to understand the practice of explanation in functional biology.

In chapter 8 I present my own account of design explanation. The basic idea of my account is that design explanations relate the way in which an organism is built, the activities of its parts, its behaviour and the condition of the environment in which it lives in terms of what is needed or useful to survive and reproduce rather than in terms of causes. There are two types of design explanation.

The first type of design explanation seeks to explain why it is useful to certain organisms to perform a certain role, for example why it is useful to Vertebrates to transport oxygen (this kind of design explanation answers a type 4a question). Such explanations proceed in the following manner:
(1) identify a need satisfied by the performance of the causal role in question;
(2) explain how that need relates to the other traits of the organism and the environment in which it lives.

The second type of design explanations seeks to explain why it is useful that a certain item or behaviour has a certain character (this type of design explantion answers a type 4b question). Such explanations proceed in the following manner:

(1) determine a causal role of the item or behavioural pattern the character of which is to be explained;

(2) explain why given this causal role and given the conditions in which this role is to be performed the role is better performed in the way it is performed than in some other way. A certain performance counts as better than another if it results in a higher fitness than that other.

I argue that such explanations are not of a causal nature. That is they do not explain how or why a certain phenomenon (transition, state or property) is brought about (neither how it is brought about in a certain organism nor how it is brought about in the course of evolution). Instead, they relate the different traits of an organism and the condition of the environment in which it lives in terms of what is needed or useful to survive and reproduce.

Chapter 2: Functions and function attributions

2.1 Introduction

In section 1 of the previous chapter I distinguished three different approaches to a philosophical analysis of function attributions. The different approaches are often seen as rival accounts of the unique meaning the notion of ‘biological function’ is assumed to have. This attitude is perhaps best illustrated by the fact that all accounts (without any exception) mention the attribution of the function to propagate the blood to the heart or the heartbeat as a typical example of a function attribution that conforms to their approach. In section 2.2 I shall show that biologists use the notion of function in a number of different ways. These different kind of functions have different explanatory roles. As will become clear in the following chapters some of the supposed rival analyses of the notion of ‘function’ pertain to different kind of function attribution. In section 2.3 I discuss the role of these different kind of function attributions in biological enquiry.

An additional aim of this chapter is to deal with some terminological issues. Unfortunately, among functional biologists as well as among philosophers there is a considerable lack of unity in the use of basic terms. In order not to get lost in a linguistic morass it is important to stick to a number of terminological conventions. I shall introduce these conventions in the course of my argument. At this point it is worth stating that I shall use the term ‘item’ to refer to any part or substance of an organism (molecules, sub cellular organelles, cells, tissues, organs, organ systems and so on). Examples of items are chlorophyll molecules, chromosomes, the cell nucleus, membranes, white blood cells, hearts, limbs, fingers, the circulatory system and so on. Biologists often use the word ‘structure’ as a countable noun to refer to parts and substances.[1] This is confusing because that term is also used as an uncountable noun to refer to the way something is built or organized (as in ‘the human heart has a four chambered structure’). I shall use the term ‘structure’ in this latter sense. Another term one often finds in the literature is ‘trait’. I shall use the term ‘trait’ to refer to: (a) the presence or absence of certain items (such as hearts and circulatory systems), behavioural patterns (such as the fanning behaviour of a stickleback) and processes (such as the beating of a heart and the circulation of the blood) of/in individual organisms; and (b) the properties (features / characteristics) of the entities under (a) (such as the structure of the heart and blood-vessels and the rate of the heartbeat) or of the

---

[1] Other terms used to refer to items: “feature” (Bock & von Wahlert 1965, Bock 1980), “element” (Dullemeijer 1974), “component” (Lauder 1986, Lauder 1990).

organism as a whole (such the size of an elephant). Roughly spoken, the term ‘trait’ refers to the presence or character of an item or behaviour in certain organisms.

2.2 Different kinds of function

I distinguish four kinds of function, namely (1) what an item or behaviour does or is capable of doing (its activities and capacities), (2) the causal role of an item or behaviour in maintaining a complex activity or capacity, (3) the survival value of a certain trait, (4) the advantages of a certain trait for which it was selected in the past. Function$_1$ attributions describe what an item does or is capable of doing, whether or not this is important in the life of the organism. Claims about the second and the third kind of function are concerned with how a certain item or behavioural pattern is important in the life of the organism. Attributions of causal roles (function$_2$) describe how a certain item or behaviour is used, that is how it contributes to a complex activity or capacity (that is to a complex function$_1$), claims about survival value (function$_3$) explain why a certain trait is useful to the organisms that have it. The fourth kind of function attribution is concerned with the past significance of a certain trait. Function$_4$ attributions sum up the effects for which that trait was selected in the past.

2.2.1 Function as activity

Functional biologists often use the term ‘function’ in opposition to ‘form’ or to ‘structure’. Roughly spoken, ‘function’ in this sense refers to what an item does or is capable of doing and ‘form’ to what the item is made of, the way it is built and the way it looks like. For example, the beating of my heart is said to be a function (i.e. an activity it performs), whereas its size, its colour and its four chambered structure are aspects of its form. I shall refer to this use of the term ‘function’ as ‘function as activity’ or ‘function$_1$’. Function$_1$ attributions describe a certain kind of observable and measurable property whether or not these properties causally contribute(d) to some complex capacity, goal, survival, reproduction, selection, evolution or whatever. Function$_1$ attributions may be qualitative (e.g. “the heart beats” or “this gland secretes mucus”) or quantitative (e.g. “the heart rate of normal human beings at rest is about 70 beats per minute, but there is much individual variation” (Keeton & Gould 1993: 850)).

The form and the function$_1$ of an item together with the environment determine what an element is capable of doing in the life of an organism. However, not everything an item does or is capable of doing is important in the life of that organism. For example, the heart’s capacity to produce sounds is of no importance in the life of most non-human organisms. Following a suggestion of Bock & Von Wahlert (1965), I shall call the activities and capacities that are of no importance in the life of an organism ‘unutilized’ and the activities and capacities that are important ‘utilized’ (Bock & Von Wahlert talk of (un)utilized functions). Bock & Von Wahlert emphasize that non-utilized activities and capacities are important in biology for two reasons. One is methodological: “we generally do not know which functions [activities and capacities] are utilized and which are not utilized by the organism” (p. 274). The second is theoretical: the utilization of a yet unutilized activity or capacity is an important mechanism of evolutionary change. An example is the utilization of the capacity of muscles to resist shearing stresses. Muscles become rigid during contraction and will support a load acting at right angles of the muscle fibres. This capacity (which is unutilized in most muscles) becomes important in the tongue flipping mechanism of frogs: in this mechanism the musculus submentalis serves as a pivot about which another muscle (the musculus geniolossus) swings (Gans 1962).

I shall refer to the form and the function$_1$ of an item together as its ‘character’. I will, occasionally, use the terms ‘functional characteristics’ (without a subscript) to refer to function$_1$ properties, and ‘structural characteristics’ to refer to form properties. This means that ‘function$_1$ attribution’ is just another word for ‘description of a functional characteristic’.

2.2.2 Function as causal role

Attributions of causal roles are concerned with the way in which a certain item or activity is important in the life of the organism. They describe how that item or activity contributes to a certain complex activity or capacity. As will become clear, attributions of causal roles are the key to understanding explanation in functional biology. For that reason I will discuss three examples in some detail.

Example 2.1: the causal role of the heart

My first example concerns a classic: William Harvey’s (1628) study of the circulation of the blood. In the beginning of the seventeenth century physiologists thought of the structure and function of the human body in terms of a tripartition. This tripartition was based upon the principal cavities they observed in dissection: the abdomen, the thorax and the head, each with their own central organ: the liver, the heart and the brain. Associated with the body cavities three fluids were distinguished: (1) venous blood containing natural spirit, formed in the liver from nutrients drawn from the stomach and transported to the organs via the veins, was supposed to supply the organs with nutrients and to remove wastes; (2) arterial blood was supposed to bring heat and life (vital spirit) to the organs; (3) animal spirit, formed in the brain and carried to the organs via the nerves, was supposed to serve the functions of movement and sensation. The heart was thought to have a central causal role in the production of heat and life. The left ventricle of the heart is the primary source of heat. Blood from the liver (were it originates) is received in the right ventricle. During contraction blood is driven from the heart to the lungs where it takes up air. Air mixed with blood is sucked back to the heart during expansion. In the left ventricle aired blood from the lungs is mixed with venous blood from the right ventricle.

Motion and heat work on it to produce the vital spirits, essential for life. Heat and life together are transported to the organs via the arteries.

Harvey's study rejects this whole picture. In the first part of *De Motu Cordis* (Chapter I–VII) Harvey improves on previous studies of pulmonary circulation. It was previously held that the active movement is expansion and that this movement originates in the chest and the lungs. Harvey argues that the heart itself is muscular and that contraction rather than dilation is the active phase of heart movement. He observes that there are no pores in the septum separating the two ventricles and draws the conclusion that all the blood must go through the lungs to get from the right to the left ventricle. The most important innovation is in the second part (chapter VII–XIV). In this part Harvey presents the idea of a continuous circulation through the whole body, from the left ventricle of the heart to the arteries, through pores in the tissues to the veins and then to the right atrium of the heart. His main argument is surprisingly simple, namely that the amount of blood pumped out of the heart during an hour greatly exceeds the weight of the organism. Harvey estimated the volume of the left ventricle as about 2 to 3 ounces. If the heart beats 65 times per minute this would amount to 10 pounds of blood in a minute or 600 pounds in an hour. This is more than three times the weight of an average man. It is more than can be supplied by the food consumed and much more than is needed for nutrition.

> It must therefore be concluded that the blood in the animal body moves around in a circle continuously, and that the function of the heart is to accomplish this by pumping. This is the only reason for the motion and the beat of the heart (Harvey 1628: 104).

Harvey's second conclusion ("the function of the heart is to propagate the blood") has become the philosopher's standard example of a function attribution in biology. It is, however, important to be aware that this attribution did not come in isolation. As is often the case in scientific research Harvey accomplished several things at once. His central claim concerns the existence of a certain complex activity (function$_1$): the circulation of the blood. This hypothesis enables him to make sense of a lot of other phenomena, such as the movement of the heart and the presence of valves in the veins. He makes sense of these items or activities by showing how they contribute to the claimed activity: the heart propagates the blood and the valves in the veins maintain the direction of the flow. This contribution to a complex activity is their causal role (function$_2$) in that system.

Example 2.2: the causal role of the thymus

My second example is the discovery of the essential causal role of the thymus in the development of the immune system in the early sixties. This example is also discussed by Canfield (1964) and by Schaffner (1993).

A standard method in discovering the causal role of an organ is to remove that organ from the body and analyze the resulting disabilities. This method had been applied innumerable times to the thymus of adult animals of different classes of Vertebrates, but it produced no results. It appeared that the thymus had no causal role whatsoever and could be removed from the body without any loss. In 1961 Jacques Miller discovered that the extirpation of the thymus in *newborn* mice does lead to serious damage. The animals suffer from atrophy of certain lymphoid organs and a shortage of lymphocytes (a certain kind of white blood cells) and die within 3 or 4 months. This is how Miller summarizes his work:

> In this laboratory, we have been interested in the role of the thymus in leukaemogenesis. During this work it has become increasingly evident that the thymus at an early stage in life plays a very important part in the development of the immunological response (Miller 1961: 748)

This discovery marked the beginning of a completely new view on the development of the immune system. Until then it was thought that all the different kinds of cells involved in the immune response were produced locally in lymphoid organs. Experiments like this one led to the hypothesis of the existence of a relatively rare population of undifferentiated stem cells. Their descendants proliferate and differentiate in the thymus into so-called ‘T-lymphocytes’. The T-lymphocytes released by the thymus colonize certain areas in other lymphoid organs, after which these organs develop and start to produce their own T-lymphocytes. The effect of thymus extirpation in adults is small because a large stock of long living T-lymphocytes has been formed at that age and because after their initial development, shortly after birth, the lymphoid organs are able to maintain themselves independent of the thymus. What the extirpation experiments show is that the thymus causally contributes to the initial differentiation of T-lymphocytes.

Canfield (1964) discusses Burnet’s (1962) overview of the state of the art. In 1962 the interpretation of the experiments was not as clear as it is now. One plausible interpretation is that

> the thymus produces and liberates into the blood the lymphocytes that pass to spleen and lymph nodes and there settle down and mature into the populations of cells that look after the integrity and security of the body (Burnet 1962: 55).

One main problem was the causal role of the lymphocytes produced by the thymus. It was found that when the thymus is removed from a particular strain of mice at birth, they do not reject skin transplants (grafts) from unrelated mice or rats (as they would normally do within 10 days). This suggests that the lymphocytes have a causal role in the recognition and removal of anomalous cells (a function “necessary for the survival and proper functioning of the individual” (p. 57)).

In chickens there appeared to be two organs involved in primary lymphocyte production: the thymus and the bursa of Fabricius. The bursa gives rise to so-called ‘B-lymphocytes’ whose descendants are responsible for antibody production, the thymus liberates those cells whose descendants are involved in the recognition of anomalous cells. Burnet believes that is “highly likely” that in mammals “the thymus also carries out the function performed by the bursa of Fabricius in the chicken” (p. 57). Later research proved this to be wrong. Initial differentiation of B-lymphocytes in mammals is now thought to take place in the bone-marrow.

As in the case of the heart the discovery of the causal role of the thymus is a complex affair. The initial experiments proved that the thymus had some causal role in the maintenance of the organism, but the nature of the activity in which the thymus takes part remained unclear. The central breakthrough came with the hypothesis of differentiating stem cells. Another important clue was the discovery that there are two lineages of differentiating lymphocytes: B-lymphocytes, involved in anti-body production, and T-lymphocytes, involved in the recognition of anomalous cells. Given this insights it is possible to make sense of the thymus by showing how it contributes to the newly discovered activity: it initiates the differentiation of T-lymphocytes.

Example 2.3: the snake’s forked tongue

My third example is a recent one: Kurt Schwenk’s explanation of “Why Snakes Have Forked Tongues” (1994). In the 1920s and ‘30s it became clear that the tongue of squamate reptiles, such as lizards and snakes, is involved in chemoreception. By flicking their tongues these organisms sample environmental chemicals and deliver it to a pair of chemoreceptors in the snout (the so-called ‘vomeronasal organ’). However, it remained unclear why the tongues of snakes (and some lineages of lizards) are forked. Until recently the functional significance of the forking was sought in the delivery phase of the tongue flicking mechanism. The idea was that the two tips are inserted directly into the two openings of the vomeronasal organ. This idea conflicts both with observations of what happens during tongue flicking and with the observation that many reptiles without forked tongues are able to deliver chemicals effectively into the vomeronasal organ.

Schwenk argues that the explanation of the forkedness must be sought in the causal role of the tongue in the sampling phase. The taxa with highly forked tongues use their tongue to follow scent trails of preys and mates. They do so by comparing the stimulus intensities of two sides of the body (a mechanism called ‘chemosensory tropotaxis’). This allows them to detect the edges of a chemical trail and follow it with minimal deviation. The larger part of Schwenk’s paper deals with the evidence for the “ineluctable” (p. 1574) conclusion that the tongue has a causal role in trail-following. This evidence consists of: (1) observations of the behaviour during trail following, (2) experiments showing that removal of the forked portion of the tongue eliminates the ability to follow scent trails, but not the delivery of stimuli to the vomeronasal organ, (3) comparative data showing a rough correlation between depth of tongue bifurcation and ability to follow chemical trails, (4) observations of the nervous system showing circuits suitable for comparing signal strength from each side of the tongue.

Once the causal role of the tongue in trail-following is established this knowledge can be used to explain the forking both at the organismal and at the historical level.

> A trail-following function for forked tongues in squamates is compelling because of its explanatory power at several hierarchical levels (Schwenk 1994: 1576).

At the organismal level the forking of the tongue is explained by observing that it meets a requirement imposed on it by chemosensory tropotaxis. In order to be able to compare stimulus intensities at two points the organism must be able to sample chemicals at two points at the same time. This is made possible by the forking:

> For a snake or a lizard to use chemosensory tropotaxis, it must be able to sense simultaneously the chemical stimulus at two points. This requirement is met admirably by the forked tongue (Schwenk 1994: 1574).

To explain the forking at the historical level Schwenk fits this trait on a phylogenetic tree. This shows that forked tongues have evolved at least twice but possibly four times. He then states (p. 1576) that “the reason for its initial evolution remains obscure”. “Probably [...] it did initially confer a performance advantage in chemical delivery to the vomeronasal organ.” Subsequent selection for increased efficiency in scent trail-following “may have caused” the tongue to become increasingly forked, a trend evident in some clades.

Characteristics of attributions of causal roles

Attributions of causal roles describe how a certain part, organ or behavioural pattern contributes to an activity or capacity of a system of which that item or behavioural pattern is a part.[2] Basically, they have the following form:

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

---

[2] The term ‘function’ is also used to refer to how a certain organ works. For example, when biologists say that the right ventricle of the heart functions like a bellow and the left one like a pressure pump they are describing the way in which the heart works. This fifth use of term ‘function’ should not be confused with function as causal role.

For example:

(1) Harvey's function attribution (the heart propagates the blood) describes how the heart (i) contributes to the organism's (s) capacity to circulate the blood (c) (namely by propagating it (f));

(2) Miller (the thymus produces T-lymphocytes) describes how the thymus (i) contributes to the organism's (s) capacity to defend the organism against invaders (c) (namely by producing T-lymphocytes (f)); and

(3) Schwenk (the tongue has a trail-following role) describes how the tongue of a snake (i) contributes to the snake's (s) capacity to find preys and mates (c) (namely by detecting the trails of preys and mates (f)).

In these examples the system s is the organism as a whole. The system might also be a part of an organism, for example the valves in the veins (i) contribute to the capacity of the veins (s) to direct the blood back to the heart (c) by preventing the blood to flow back (f). Attributions of causal roles may simply state that an item contributes to a certain complex activity (e.g. "the heart has a function in circulation", "the thymus plays a part in the development of the immunological response") or describe that causal role in more or less detail (such as in "the heart is the source of energy of the circulatory system", "the thymus initiates the differentiation of T-lymphocytes", "the tongue has a trail-following role"). Sometimes an attribution of a causal role is even more complex and describes the activity (function$_1$) by means of which the causal role is performed in addition to the causal role itself. An example would be the statement "the heart contributes to circulation by beating".

In the case of the heart, a newly discovered activity (function$_1$), the circulation of the blood, resulted in a change in the causal role (function$_2$) attributed to the heart. Insight in the causal role of the thymus came only after the discovery of a yet unknown activity (function$_1$) to which the thymus contributes: the process of differentiating lymphocytes. In the case of the forked tongue no new activities had to be discovered: it turned out that the tongue has a causal role in a known activity: trail-following.

The distinction between function as activity (function$_1$) and function as causal role (function$_2$) is a distinction between statements describing a certain kind of observable properties and statements describing how (the activity of) that item contributes to some complex activity or capacity. If one reports about a certain item without taking into account the effects of that items on its environment (e.g. "the heart beats") one describes the activity of that item. Attributions of causal roles tell us about the effects of (the activity of) certain items in a larger context[3]. It depends, of course, on the environment whether or not the heart beats and how fast it will beat but in saying that the heart beats one does not say much about the influence of the heart on the

---

[3] And so do attributions of survival value (function$_3$) (as will become clear in the next section).

organism. However, if one states that “the heart propagates the blood” one takes into account a larger context of the heart’s activity (namely the circulatory system), which makes it appropriate to speak of the causal role of that item (in that larger context). Similarly, if a certain behaviour is described as “swimming” or “walking” or “flying” one does not take into account the influence of that behaviour on the environment, which means that one describes an activity. However, if one says that walking serves to acquire food one details the causal role of that behaviour in a larger context.[4]

2.2.3 Function as survival value

The term ‘function’ is also used to refer to the selective advantages of a certain trait in comparison with another trait. I call this use of the term ‘function’ ‘function as survival value’ (function$_3$). The study of survival value aims to find out why it is useful that a certain item or behaviour is present or absent and/or why it is useful to certain organisms that a certain item or behaviour has a certain character. Such an account is seldom given in one sentence.

Example 3.1: the survival value of egg shell removal in birds

An example of research into survival value is the study of the egg shell removal behaviour in birds. Many birds remove the empty egg shell after the chick has hatched. In most species, the egg is picked up and dropped at some distance of the nest, but there are several other ways to dispose of an empty egg shell. Tinbergen and his students (1962) performed a series of now classical experiments on this pattern of behaviour in the black-headed gull. Their study concerns the causes of the egg shell removal behaviour (i.e. the stimuli that elicit this behaviour) as well as its survival value. I restrict my account to the latter part of the study.

Tinbergen c.s. list a number of different ways in which the presence of an empty egg shell might be disadvantageous to the brood or to the parents. It might be that the sharp edges of the empty shells would injure the chicks, that the empty shells would interfere in some way or other with brooding, that the empty shells would provide a breeding ground for bacteria and moulds, or that the shells would draw the attention of predators. The egg removal behaviour is not performed by the kittiwake and the sandwich tern. These birds live in conditions were predation is low. This suggest that the main function of the behaviour is the maintenance of the camouflage of the brood (in all other cases there is no reason why the kittiwake and the sandwich tern should lack the response).

---

[4] Whether or not a certain kind of movement counts as swimming, walking or flying depends, of course, on the environment, but, in labelling a certain behaviour as walking one does not imply much about the effect of that behaviour on the environment.

The hypothesis that the egg shell removal behaviour serves to maintain the camouflage of the brood presupposes that the brood *is* camouflaged. For that reason, Tinbergen c.s. performed a experiments which test this presupposition. These experiments show that carrion crows, herring gulls and black-headed gulls find eggs that are painted white more easily than normal eggs. It is concluded that

> 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; in other words their colour acts as camouflage (Tinbergen et al. 1962: 80/81).

The second series of experiments shows that predators find normal eggs covered with some straws of grass more easily if those eggs are given an egg shell at 15 cm. distance. It is concluded that

> the near presence of an egg shell helps Carrion Crows and Herring Gulls in finding a more or less concealed, camouflaged prey, and that therefore egg shells would endanger the brood if they were not carried away (Tinbergen et al. 1962: 82).

In a third series of experiments it is shown that this “betrayal effect” decreases rapidly with an increasing distance between eggs and shells. It is concluded the egg shell removal behaviour has survival value because it helps to keep predators away:

> The conclusion of this part of the study must therefore be that the eggs of the Black-headed Gulls are subject to predation; that in tests outside the colony the number of eggs found by Carrion Crows and Herring Gull is lower than it would be if the eggs where white; that the proximity of the egg shell endangers the brood; and that this effect decreases with increasing distance. While it will now be worth investigating the predators’ responsiveness to eggs and shells in more detail, the facts reported leave little room for doubt about the survival value of egg shell removal as an antipredator device. Whether or not the response has other functions is of course left undecided (Tinbergen et al. 1962: 85).

In addition, Tinbergen c.s. observe that black-headed gulls wait for an hour or two before removing the empty egg shell, whereas waders such as ringed plovers and oystercatchers remove the empty shell immediately. They point out that this lack of promptness (of the removal action in black-headed gulls) has survival value because it tends to reduce predation by other black-headed gulls. After hatching, it takes the chick a few hours to dry. In any colony of black-headed gulls there are some individuals who prey selectively on nearly hatched eggs and wet chicks. They take the chick when the parents leave the nest to attack other predators or to remove the egg shell. Dry chicks are left alone. Waders don’t live in colonies and their bills aren’t fit for eating chicks. Tinbergen c.s. draw the following conclusion:

> We feel justified therefore to ascribe the lack of promptness of the response to this tendency of some members of the colony to prey on wet chick (Tinbergen et al. 1962: 110)

In the discussion section of the paper the authors discuss the anti-predator system of black-headed gulls as a whole. They explain among others that the camouflage of the eggs is important because breeding gulls leave the brood at the first sign of danger. This behaviour in turn is related to the suspicious colour of the adult: in species in which the breeding adult itself is camouflaged (such as ducks and pheasants) it usually remains on the nest. Tinbergen c.s. also show that the total system is best explained as a compromise between conflicting demands. An example of such a conflict was mentioned above: removal of the egg shell might help to defend the brood against carrion crows, but it increases the chance that the chick is eaten by a neighbour. A delayed response seems the best compromise between these conflicting demands.

Example 3.2: social foraging by black-headed gulls

Bleak-headed gulls usually fish in groups. Two hypotheses have been advanced concerning the survival value of this behaviour: the parasitism hypothesis and the co-operative hypothesis. According to the parasitism hypothesis (Ward & Zahavi 1973) the flocks consists mainly of birds that have followed another gull (“the leader”) who has found a rich fishing place. The behaviour is advantageous to the followers because it enables them to locate food resources they would not have found otherwise. To the leader the following behaviour is neutral or negative (because it depletes the food source found by the leader). A group of Swedish researchers argues for the co-operative hypothesis (Gotmark, Winkler & Anderson 1986). On this hypothesis the flocking behaviour is useful because it enables each gull to catch more fish than it would have done if it fished alone (a fish that tries to escape from one gull might run into the beak of another). The research group put different numbers of gulls in a large aviary with a fishing pool and counted the number of fish the gulls catched. They found that birds in groups catch more fish (per gull) than do solitary birds in the same time and that the number of fish catched per gull increases if the group size increase. They draw the conclusion that flock feeding has survival value because it allows all gulls to catch more fish than they would have done otherwise.

Example 3.3: behavioural differences between two species of gazelles

In an issue of the *Journal of Zoology* K. Habibi, C.R Thouless & N. Lindsay (1993) compare the behaviour of the two species of gazelles that live in Saudi Arabia: the sand gazelle (*Gazella subgutturosa*) and the mountain gazelle (*Gazella gazella*). Their study reveals differences in group size and composition, mating season, territory size, territory marking, sexual and agonistic behaviour. Sand gazelles are seasonal breeders that form large herds of up to 65 individuals. Outside the breeding season, these groups consist of individuals of both sexes. During the breeding season, the males round up the females in harems (consisting of 15–20 individuals). The males hold territories which they defend vigorously, often engaging in prolonged fights. The territories are considerably smaller than those of the mountain gazelle and are marked both by urination marks and by a substance secreted by the preorbital glands. Mountain gazelles reproduce during the whole year. They live alone or in small groups of 2–5 individuals (usually consisting of a female and her offspring). Mountain gazelle males hold territories during the whole year. They exhibit threat displays rather than engaging in actual combat. The territories are about twice as large as those of the sand gazelle and they are marked by dung piles deposited at the border of the territory. Instead of chasing and herding all females that enter his territory (as the male sand gazelles do) a mountain gazelle male approaches a female after she has entered his territory and checks if she is in oestrus. If she is, he will closely guard her, attempting to prevent that she leaves his territory.

In the discussion section of their paper the authors try to relate the differences in terms of survival value. Their account is rather speculative, but it gives a good insight in the way in which appeals to survival value are explanatory.

The authors argue that the differences in territory behaviour (harems / individual territories) and the differences in the duration of the territory holding “may be attributable to” the difference in the timing of reproduction: the survival value of herding females in harems is greater in the case of seasonal breeding whereas the survival value of holding individual territories is greater in the case of reproduction during the whole year. During the rutting season of sand gazelles, a large proportion of adult females will be either sexually receptive or about to become receptive. Therefore, it is worthwhile for a male sand gazelle to round up all females, regardless of their immediate status and keep them in his territory. In contrast, the probability that a particular female mountain gazelle will soon become receptive is fairly low, at any time of year. Therefore, for a male mountain gazelle it is of no use to herd all females that come across his path. Rather, it is worthwhile to keep an individual territory during the whole year as this increases his chances of meeting receptive females.

The differences in the duration of the territory holding in their turn explain the differences in agonistic behaviour (fight / threat): the male mountain is familiar with his territory and with his neighbours which will increase the survival value of threat displays over actual combat (an invader has the disadvantage of fighting on unknown land, neighbours know what the other is worth).

Next, the authors suggest that the difference in territory marking “may be connected with” the difference in territory size: “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”.

The difference in the timing of reproduction is somewhat puzzling (according to the authors). Seasonal breeding is usually related to seasonal variation in food abundance. Northern populations of both species face this condition, however, only one of the species is a seasonal breeder. Perhaps the difference is related to a difference in “life strategy”: sand gazelles live in open country and travel over long distance to find their food; mountain gazelles, on the other hand, live in more broken areas and stay in the same place. “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”.

The snake’s forked tongue, revisited

The examples above are concerned with ethology. The study of survival value is as important in functional morphology, as it is in ethology. In fact, I have already discussed an example of an account of survival value in functional morphology, namely in example 2.3 (the snake’s forked tongue). As I said there, after having established that the tongues of snakes have a causal role in trail-following, namely to sample chemicals at two points at the same time, Schwenk observes that the tongue is able to perform this causal role if it is forked but not if it is blunt. This means that the forked character of the tongue is useful to the organisms that have it because it enables the tongue to perform its causal role to sample chemicals at two points simultaneously.

The thymus, continued

Note that, the story of the thymus, as I presented it above (example 2.2), ends with the attribution of a causal role, not with a claim about survival value. The research in the early 1960s made plain the causal role of the thymus (namely to initiate the differentiation of T-lymphocytes) but it remained unclear why it is useful to land vertebrates to have a special organ to perform this causal role. It also remained unclear why it is useful to have a complex mechanism in which differentiation is initiated in the thymus and continued in the lymphoid organs.

The first attempts to answer such questions date from the late 1980s. Rodney Langman’s *The Immune System* (1989) is an excellent example of an explanation in terms of survival value. Langman seeks to explain the way in which the immune system is organized. He does so by showing that the way in which the immune system is actually organized is more useful to the organisms that have it than other conceivable ways of organizing this system.

Unfortunately, the explanation is too complex to summarize it adequately in a sentence or two. The main line of explanation is this: in order to avoid self-destruction the immune system of an organism must be able to distinguish between what belongs to the organism (“self”) and what belongs to an invader (“non-self”). The knowledge of which things count as self and which count as non-self might be genetically fixed or it might be learned in the course of ontogenetic development. Langman argues that, in the conditions in which land vertebrates live, the latter mechanisms is to be preferred (if the criterion by which the immune system recognizes self is genetically fixed, invader cells might learn what the criterion is and use this knowledge to present themselves as self).

He also argues that the best mechanism to learn to distinguish self from non-self uses the following distinction between self and non-self: self is what is present during a long period, non-self is what is present during a relative short period. Learning to distinguish between self and non-self is a complex process. It is therefore more efficient to delegate the task to recognize non-self to specialized cells (the T-helper cells) which regulate the other immune cells (this saves the costs of learning all immune cells to distinguish between self and non-self).

Furthermore, as the learning process is complex it is more efficient to locate this process in a specialized organ and transport the ability to distinguish between self and non-self afterwards to the places where this ability is needed than to generate this ability everywhere where this ability might be useful. This organ is the thymus. Initially, young organisms produce many different types of proto-T-lymphocytes, each type is able to recognize another antigen (but it is not yet able to activate other immune cells). What happens in the thymus is that those proto-T-lymphocytes which recognize antigens of parts which are present during a long time are deactivated, whereas the other types of T-lymphocytes acquire the ability to activate other immune cells. The initialized T-lymphocytes are transported to the lymphoid organs were they mature and proliferate.

In sum, Langman’s main point is that in the conditions that apply to land vertebrates it has survival value to have a specialized organ to initiate the differentation of T-lymphocytes because this allows for a more efficient mechanism to distinguish self from non-self than distributed production would do.

Characteristics of claims about survival value

Claims about survival value state why under certain conditions it is more useful for an organism to have a trait it has rather than another one. Such claims have the following basic form:

> trait *t* has survival value in comparison with trait t' under conditions *c* because of ... (follows an explication of why an organism in condition *c* would be worse off if it had t' instead of *t*).

Examples are:

“if the eggs are subject to predation by carrion crows and herring gulls (*c*) it is more useful to remove the empty shell after the chick has hatched (*t*) than to leave it near the nest (t') because the empty egg shells would break the camouflage of the remaining eggs”,

“flock-feeding (*t*) is more advantageous to gulls than solitary feeding (t') because it enables them to locate food resources they would not have found otherwise“,

“flock feeding (*t*) has survival value (in comparison to solitary feeding (t')) to gulls because it allows all individuals in the flock to catch more fish than they would have done otherwise”, "if the territory is large (*c*) dung piles (*t*) are better territory markers than scent-marks (t'), because the latter are more volatile and it would not be possible to replenish them fast enough",

"a forked tongue (*t*) is more useful than a blunt one (t') to organisms that use their tongue in chemosensory tropotaxis (*c*) because a forked tongue enables them to sample chemicals simultaneously at two places, which would not be possible with a blunt one".

The explication might be quite a complex piece of reasoning, as we have seen in the case of territory behaviour and also in the case of the immune system.

Claims about survival value are essentially comparative: the presence of a certain item or behavioural pattern or the character of such an item or behaviour has survival value in comparison to another trait. Typically the comparison is counterfactual: an organism with a certain trait *t* is compared to a hypothetical organism that is similar to the real organism except that *t* is lacking or replaced by another trait t'. Arguments concerning survival value make ample use of a kind of counterfactual which I shall call 'functional counterfactual'. Functional counterfactuals state that an organisms would have certain disadvantages if a trait it has would be replaced by another one (or if they would lack that trait). Examples can be found in the examples above: "the natural colour of the eggs makes them less vulnerable to attack than they would be if they were white", "the eggs shells would endanger the brood if they were not carried away", 'flock feeding enables gulls to locate food resources they would not have found otherwise', 'flock feeding enables each gull to catch more fish than it would have done if it fished alone', 'it would not be possible for a male to replenish scent marks fast enough to be effective in a large territory', 'neonates and lactating females would be under stress when travelling long distances to new feeding grounds'.

The conditions in which the trait in question (*t*) is more useful than the alternative trait (t') can be properties of the organisms to which the attribution applies (internal conditions) and / or characteristics of the environment in which those organisms live (external conditions). Quite often the conditions are not completely specified. In the case of flock feeding for example, the conditions remain vague. It is, however, clear that the behaviour of the prey (fish) is one on the conditions that makes flock-feeding useful. It is one of the aims of research in functional biology to identify the relevant conditions accurately.[5] It is important not to confuse the conditions that trigger a certain reaction in an organism and the conditions in which this response is useful. For example, the shadow of a bird of prey may cause a hiding response in a hare, but this behaviour is useful only if there really is a bird of prey. So, the shadow is a condition which

---

[5] In attributions of survival value names of taxa (such as gulls, gazelles, snakes) or pseudo-taxa (such as fishes) usually refer to all individuals that satisfy a vague set of properties rather than to lineages of individuals.

triggers the behaviour and the presence of a bird of prey is the condition in which this behaviour is useful. Only the latter kind of conditions occur in attributions of survival value.[6]

The survival value of the presence or character of an item or behavioural pattern is typically assessed in relation to the causal role of that item or behaviour pattern in the maintenance of the organisms in study. For example, the survival value of feeding in flocks (rather than on your own) is that the causal role of that behaviour (feeding) is performed better (more fish are caught) in the case of flock feeding than in the case of solitary feeding; the survival value of dropping dung piles rather than scent marks is that in large territories dropping dung piles performs its causal role (territory marking) better than leaving scent marks; the survival value of the tongue being forked is that it enables the tongue to perform its sampling role and so on. This means that attributions of survival value (function$_3$) typically depend on a preceding attribution of a causal role (function$_2$).

The ultimate criterion for determining which trait is better is the fitness of the organisms that have those traits. ‘Fitness’ is a technical term of evolutionary theory. It is best interpreted as the propensity to survive and reproduce in a particular environment (Mills & Beaty 1979). More specifically the fitness of an organism in a specified type of environment is defined as the expected number of offspring of that organism in that type of environment. Because of its dispositional nature fitness can not be measured directly. Evidence for judgements about fitness differences comes from two sources. First, judgements about fitness differences might be based on measurements of the actual number of offspring. If in a certain environment individuals with a certain trait t have on average more offspring than individuals with an alternate trait t' this is evidence that in that environment individuals having t are fitter than individuals having t'. Second, judgements about fitness differences might be based on an examination of physical design. For example, if eggs are subject to predation, it is plausible to assume that the fitness of the parents increases if the eggs are better camouflaged. To determine fitness on the basis of physical design several criteria are used (such as the number of fish catched in a certain time, the number of females fertilized, the efficiency with which a territory is hold). These criteria usually come down to the efficiency with which a certain causal role is performed. It is assumed that these criteria correlate with fitness but the exact relation between the criteria and fitness is often left unspecified. Attributions of survival value typically state that under certain conditions (e.g. large territories) a certain task (e.g. territory marking) is performed more efficiently in one way (e.g. by means of dung piles) than in another (e.g. by means of scent marks). It is assumed that individuals that perform territory marking more efficiently have a

---

[6] Biologists often use the terms “proximate cause” or “immediate cause” for the conditions that trigger a certain response and “ultimate cause” for the conditions in which a certain response is useful. This terminology is misleading because the so-called “ultimate causes” are not causes at all: they do not bring about the response.

higher fitness than individuals that perform territory marking less efficiently but the exact connection between fitness and efficient performance is left open.

Because causal roles and survival value are so often confused I will give an elaborate account of the differences in section 6.2. For the moment it suffices to mention the main differences. Causal roles are attributed to items or behavioural patterns, whereas claims about survival value concern traits, that is they concern the presence or character of an item or behavioural pattern. For example, one speaks of the causal role of the tongue and of the survival value of the tongue being forked (rather than blunt). An important difference between function as causal role (function$_2$) and function as survival value (function$_3$) is the kind of context which is taken into account. Both attributions of causal roles (function$_2$) and attributions of survival value (function$_3$) are concerned with how a certain item or behavioural pattern is important in the life of an organism. Attributions of causal roles take into account how an item is used to achieve some complex activity or capacity. Attributions of survival value take into account how the presence or character or activity of an item or behaviour influences the life chances of an organism and/or its descendants. One may speak of the causal role of an item or behaviour *in* a certain organism and of the survival value of that item or behaviour having a certain character *for* an organism in a certain environment. Another important difference is this: attributions of survival value are comparative and relative to a certain environment; attributions of causal role are not comparative and independent of the environment.

2.2.4 Function as selected effect

Evolutionary biologists sometimes use the term ‘function’ in a historical sense to refer to the effects for which a certain trait was selected in the past. I shall call this use of the term ‘function’ ‘function as selected effect’ or ‘function$_4$’. Functions in this sense are singled out by their role in the evolutionary history of the trait to which the function is attributed. Selected effects are past effects that help to explain the trait’s current presence and/or frequency in the population.

Example 4.1: the function of inquilinism

A clear example of this use of the term ‘function’ can be found in the following quote from George Williams’s classic on “Adaptation and Natural Selection” (1966) (this passage is also quoted by Wright 1973: 92/3):

> I remember a particularly relevant oral discussion of the function of inquilinism among pearlfishes. These slender-bodied fishes live in the respiratory systems of sea cucumbers. They apparently emerge at night to forage, and return at dawn to their hosts. They are largely without pigment, and there is some evidence that they are harmed by exposure to daylight. The question arose: Do these fishes enter the sea cucumber

> to avoid light, or do they do so to avoid predators? The feeling of the group seemed to be that if the behavior fulfills both needs, it must be regarded as having a dual function. This is a physiological valid conclusion, but teleonomically naïve. The two needs are surely not historically coordinate. All fish are under pressure to avoid predators, but very few are damaged by exposure to light. This must have been the condition of the ancestors of pearlfishes. The habit of entering holothurians developed as a defense against predators, and the fish became extremely specialized in behavior and physiology for exploiting the advantages of inquilinism. This required or permitted the degeneration of a number of adaptations: the caudal fin disappeared; the eyes were reduced; an the integumentary pigments and other defenses against light were reduced in effectiveness. In this way inquilinism became a necessary part of the defense against physiological damage by light. It was not, however, as a defense against radiation that the behavior originated (Williams 1966: 266/7).

In this quote Williams makes a distinction between the needs a trait satisfies and the functions it has: the behaviour of inhabiting sea cucumbers (holothurians) during the day satisfies both the need to avoid light and the need to avoid predators, but only the latter effect counts as a function of that behaviour. To count as a function the effect must not merely satisfy a need but the behaviour that satisfies the need must have *evolved as a means* to do so. The habit to live in sea cucumbers evolved as a means to avoid predators and so its function is to avoid predators. Once this habit had evolved the pearlfishes became dependent on living in sea cucumbers not only to avoid predators but also to avoid light. Hence, the behaviour satisfies the need to avoid light. However, the behaviour did not evolve as a means to avoid light and, according to Williams, avoiding light should, therefore, not count as a function of the behaviour. Williams calls the effects of a certain trait on which the organisms having the trait became dependent after the initial evolution, of that trait "secondary needs".

Characteristics of attributions of selected effects

Williams does not expound on what it means to say that the habit of entering sea cucumbers "developed as a against defense predators". The idea of 'a trait having evolved as a means to some effect' can be unpacked with help of a schematic account of the evolution of that trait. Suppose the habit of entering sea cucumbers emerged in the following way. Imagine an ancestral population of proto-pearlfishes that have not yet evolved the habit of entering sea cucumbers. In this population a mutant pearlfish developed the habit of entering sea cucumbers. As a result of this habit mutant pearlfishes were less easily catched by predators than were the original variants. As a result mutant pearlfishes produced more offspring than did the original variants. This in turn effected a rapid spread of the mutation, up to the point that only inquilinistic pearlfishes remained. According to this scenario, the habit of entering sea cucumbers evolved because it enabled its bearers in an ancestral population to avoid predators more effectively than its rivals that did not perform this behaviour. This might be abbreviated by saying that the habit of entering sea cucumbers evolved ‘as a means to’ avoid predators or by saying that the trait evolved ‘as the result of selection for’ avoiding predators. More generally, a trait evolved ‘as a means to’ a certain effect or ‘as the result of selection for’ that effect if the trait evolved because that effect conferred a greater relative fitness to its bearers in an ancestral population. This effect is its function as selected effect (function$_4$). Hence, a function as selected effect (function$_4$) is an advantage a certain trait conferred to its bearers in an ancestral populations that gave rise to the subsequent evolution of that trait (that is of the increase of organism having that trait).

To avoid misunderstandings it should be noted that the selected effects of a certain trait are not necessarily a subset of the effects of that trait that have survival value. Functions as selected effects concern past effects that need not occur in the current organisms having that trait. For example, if pearlfishes and sea cucumbers are kept in an aquarium without predators the inquilinistic behaviour does not have avoiding predators as survival value but that effect is still the effect for which the behaviour evolved. Hence, a clear distinction should be made between claims about survival value (function$_3$) and attributions of selected effects (function$_4$). Functions as selected effect are singled out by their role in the evolutionary history of the trait that has that function; functions as survival value are singled out by their value for the organisms that have that trait. An attribution of survival value is a claim about the importance of a certain trait for the survival, reproduction and/or fitness of the organisms that have that trait. An attribution of a selected effect is a claim about the evolutionary history of that trait. Claims about survival value tell you how a trait *could be* important for the *organisms* that have it, attributions of selected effects tell you what *has been* important in the evolution of the *trait* to which the function is attributed.

2.3 The role of function attributions in biological enquiry

2.3.1 Introduction

In the preceding section I distinguished four kinds of function attributions in connection with four kinds of function. In this section I discuss the role of these different kinds of function attributions in biological enquiry. I restrict myself mainly to the parts of biology that are called ‘functional animal morphology’ and ‘ethology’ (because these are the disciplines with which I am best acquainted). In order to specify the role of the different kinds of function attributions I distinguish seven types of questions that guide research in both these disciplines. I detail the role of the different kind of function attributions by stating how those different kind of function attributions are used in posing and answering the seven different types of questions. These questions concern: (1) the character of an item or behaviour, (2) its causal roles, (3) the causes and underlying mechanisms, (4a) the survival value of performing certain tasks, (4b) the survival value of having a certain character, (5) ontogeny, and (6) evolution. The different types of questions and the types of answers produced to these questions are summarized in table 2.1. In section 2.3.2 I present the different types of questions and answers in general. In section 2.3.3 and 2.3.4 I show by means of examples that this frame work of questions applies to functional morphology respectively ethology. My examples concern the heart of mammals and the singing behaviour of birds. In section 2.3.5 I specify the role of the different kinds of function attributions by relating these attributions to the framework.

Table 2.1: Different issues concerning the form and function of a certain item or behaviour

Problem area	**Typical questions**	**Type of answer**
(1) character	what does it look like? how is it built? what is its structure?	description of the form of an item or behaviour
	what does it do? what is it capable of doing?	description of the function$_1$ (activity) characteristics of an item or behaviour
(2) causal role	how is it used?	attribution of one or more causal roles
(3) causes and underlying mechanisms	how does it work?	physiological explanation
(4a) survival value of performing certain tasks	why does the organism have an item / behaviour that performs this role?	design explanation (of the need to perform a certain causal role)
(4b) survival value of having a certain character	why is it built the way it is? why does it work the way it does?	design explanation (of the character of an item or behaviour)
(5) ontogeny	how did it develop in the course of the ontogeny?	developmental explanation
(6) evolution	how and why did it evolve?	evolutionary explanation

2.3.2 Seven types of questions and their answers

Introduction

When functional morphologists and ethologists study a certain item or behaviour they typically ask seven types of questions. These questions concern:

1) the character of the item or behaviour in question,
2) the causal roles of the item or behaviour in question,
3) the causes and underlying mechanisms resulting in the realization of those causal roles,
4a) the survival value of performing the causal roles the item or behaviour in question performs,
4b) the survival value of having the character that the item or behaviour in question has,
5) the development of the item or behaviour in question in the course of the ontogeny,
6) the development of the item or behaviour in question in the course of evolution.

Functional morphologist ask these questions typically about an item under study (such as the heart), ethologists ask these questions typically about a behaviour (such as a bird's song).

The first type of questions concerns the character (form and function$_1$) of the item or behaviour under study. What does the item look like? How is the item built? What does it do? What is the structure of the behaviour? An example of a question of this kind in morphology is the question 'how is the heart built?'; an example from ethology is 'what is the structure of a bird's song?'. Research into this kind of question aims for accurate descriptions of the item or behaviour under study.

The second type of questions concerns the way in which the item or behaviour under study is used by the organism. Examples of questions of this type are 'what is the causal role of the heart?' and 'what functions does a bird's song have?'. These questions are answered by means of one or more attributions of a causal role (function$_2$ attributions). Examples are the attribution of the causal role to pump the blood around to the heart and of the causal role to claim a territory to bird's songs. Note, that causal roles are attributed to items or behaviours, not to their character.

Questions of the third type ask 'how does the item or behaviour in question work?'. That is, how is that item or behaviour able to perform the causal roles attributed to it in answer to a type (2) question? Examples are 'how is the heart able to pump blood?' and 'how are bird's songs produced?'. These questions concern the causes and underlying mechanisms of the activity of an item or of the behaviour of an organism. An answer to such a question is usually called a 'causal explanation' by biologists. As I use the term 'causal explanation' in a broader sense, I shall use the term 'physiological explanation' to refer to explanations in this area of research. Physiological explanations come in (at least) two different kinds. The first kind of physiological explanation explains certain changes in the state of an organism (such as changes in the frequency of the heartbeat, or changes in a bird's readiness to sing) as the effect of preceding changes in the organism or its environment. For example, changes in the frequency of the heartbeat are explained by changes in the activity of the nerves that innervate the heart, which in turn are explained by, say, the fact that the organism hears the alarming call of another organism. The second kind of physiological explanation explains the properties of an organism (including its capacities) as the result of underlying structures and mechanisms. An example would be an explanation of skin colour in terms of cellular pigments and their arrangement. An important subtype of the second kind of physiological explanation are the kind of explanations which I shall call ‘capacity explanations’. A capacity explanation explains a capacity of an item of an organism (or of the organism as a whole) by appeal to the capacities of the parts of that item (or of the organism as a whole) to perform a series of tasks which add up to the capacity to be explained. For example an explanation of an organism’s capacity to circulate oxygen would point out that oxygen circulation is brought about by a system of vessels which contain blood. The blood carries the oxygen and is pumped around by a heart. The two kinds of physiological explanations are related in the following way: explanations of the second type are concerned with the mechanisms that connect the causes and effects mentioned in explanations of the first type. For example, an explanation of the second type might concern the mechanism that bring about changes in the frequency of the heartbeat in response to changes in the activity of the nerves that innervate the heart..

The next two types of questions are both concerned with survival value. For that reason I have labelled them (4a) and (4b). Research into these questions aims for an explanation of the way in which an organism is built and the way in which it behaves in terms of the utility of that design in the environment in which that organism lives. Such explanations are usually called ‘functional explanations’ by those who engage in them. Because many philosophers use the term ‘functional explanation’ in a different sense I shall use the term ‘design explanation’ to refer to this kind of explanation. Design explanations are used both (in answer to type (4b) questions) to explain why it is useful to certain organisms that a certain item or behaviour has a certain character (for example why it is useful to snakes that their tongues are forked), and (in answer to (4a) questions) to explain why it is useful to certain organisms to have an item which performs a certain causal role (for example why it is useful to vertebrates to have a system which circulates oxygen). This utility is related to the other traits of those organisms and the state of the environment in which they live. A design explanation is a claim about the survival value of performing a certain role or of having a certain character.

Questions of type (4a) ask ‘why is it useful to the organism to have an item or behaviour that performs the causal roles attributed to the item or behaviour under study?’. Examples of such questions are ‘why is it useful to circulate the blood?’ and ‘why is it useful to defend a territory?’. Design explanations that answer such questions identify a need that is satisfied by the performance of the role in question and relate that need to the other traits of the organism and to state of the environment in which it lives. An example is the explanation of the presence of a circulatory system in vertebrates by pointing out that because of their size these organisms need to transport oxygen actively rather than passively (that is by mere diffusion). It is a law of physical chemistry that the rate of diffusion decreases proportionally to the distance over which diffusion takes place. In large organism the distance between the inner organs and the outside is such that the rate of diffusion is too slow to get enough oxygen to the organs. This problem is solved by transporting oxygen actively from the outside to the organs. I discuss this example in more detail in section 4.2.3.

Questions of type (4b) ask ‘why it is useful that the item or behaviour under study has the character it has?’. Examples of such questions are ‘why is it useful that the heart of mammals consists of four chambers?’ and ‘why is it useful to defend a territory by singing (rather than by attacking intruders)?’. Design explanations that answer such questions proceed by showing that (in the conditions applying to the organisms concerned) the causal role in question is better performed by an item or behaviour with the character to be explained than by items or behaviours with some plausible alternative character. An example is the explanation of the forkedness of the snake’s tongue (example 2.2 of section 2.2.2) by pointing out that this form meets a requirement imposed on it by the trail-following role of the tongue, namely the requirement to sense simultaneously the chemical stimuli at two points.

The next kind of questions (5) consists of questions that concern the ontogenetic development of the item or behaviour under study. How did this item or behaviour develop in the course of the ontogeny and how is this development controlled? Examples of such questions are ‘how does the heart develop and how is this development regulated?’, ‘how do bird songs develop?’, ‘is the song pattern innate or learned from parents?. The explanations proposed in answer to questions of this kind are usually called ‘developmental explanations’. Developmental explanations explain how a certain trait arises in the course of the ontogeny. In the example of the circulatory system, a developmental explanation would (among other things) point out that the initial differentiation of blood vessels is probably caused by a process of induction (see for example Balinsky 1975: 410). The first blood vessels develop before circulation starts. If the heart rudiment is removed before it starts to beat the large blood vessels continue to develop for some time. Further development depends on the direction and amount of the blood flow through these vessels. Developmental explanations and physiological explanations shade into each other. The main difference is that developmental explanations are concerned with transitions that usually occur only once in the lifetime of an organisms and physiological explanations with transitions that may occur repeatedly.

The part of biology that is concerned with individual development is called ‘developmental biology’. Developmental biology is not limited to type (5) questions about the traits of the adult or larva. Indeed, developmental biology addresses questions of all seven types, but they are asked in regard to development rather than in regard to the adult individual. For example, most textbooks on developmental biology address the following type (4b) question: ‘why does the heart starts to beat early in development (much earlier than all other organs start to perform their causal role)?’. The answer to this question (the embryo at this stage already needs a system to transport oxygen) is a design explanation. Another example concern the items, activities or structures which develop in an embryo and which are not precursors of the organs etc. of the adult or larva but have a causal role in the maintenance of the embryo. Examples are organs which store and utilize yolk (the yolk sac and vitelline blood-vessels), organs which store wastes (such as the allantois) and organs which protect the embryo from desiccation (such as the amnion, the chorion and the cavity between them). About such organs all seven questions are asked. I will not pay much attention to developmental biology in the remainder of this book.

The last kind of questions (6) consists of questions concerning the evolution of the item or behaviour under study. How and why did this item or behaviour evolve and how and why did it acquire the character it has? Examples of such questions are ‘how did the heart became four chambered?’ and ‘how did bird songs became complex?’. Explanations that answer questions of this kind are called ‘evolutionary explanations’. Evolutionary explanations explain how a certain trait developed in the course of the history of the lineage. There are several processes that may explain how form, function$_1$ and behaviour changed in the course of the history. Not all these processes are evolutionary processes. For example, humans are larger now than in the past. This change is a direct effect of better nutrition and medicine and does not involve genetic change. For that reason, this historical change is not evolutionary. Evolutionary processes include mutation, gene flow, recombination, selection and genetic drift. I shall call evolutionary explanations that focus on evolution by natural selection ‘evolutionary selection explanations’.[7] Evolutionary selection explanations explain the presence or character of a certain item or behaviour by telling how and why natural selection modified that item or behaviour in the course of the history.

The part of biology that is concerned with evolution is called ‘evolutionary biology’. Evolutionary biology is broadly defined as the study of the history and mechanisms of evolution. The two main kinds of evolutionary phenomena are branching and character change. Branching is the splitting of lineages of populations into separate branches. Character change is the change of frequencies of characters of individuals in a population over generations.

In the remainder of this section I will give some examples of the application of these types of questions in functional morphology and ethology.

---

[7] In chapter 7 I distinguish a second kind of selection explanation, namely equilibrium selection explanation.

Functional morphology

Functional morphology is usually broadly defined as the study of the relations between form and function.[8] Its main concerns are:

1) to describe the forms and the functions$_1$ (activities) of the parts and organs of the body (and of the body as a whole);

2) to find out what the causal role is of those items in the maintenance of the organism;

3) to explain the character and capacities of those items and the changes in their state in terms of underlying structures and mechanisms;

4a) to explain why it useful to perform the causal roles those items perform in the context of the organism as a whole and the environment in which it lives;

4b) to explain the character and capacities of those items in terms of their utility in the context of the organism as a whole and the environment in which it lives;

5) to explain how those items (their form, activity and causal role) have developed in the course of the ontogeny;

6) to explain how those items (their form, activity and causal role) have evolved in the course of history.

Consider, for example, how these questions apply to the heart. The discussion of “the structure and function of the heart” in an arbitrary text on functional morphology (for example Johansen 1977: 387-389) starts with a description of the structure and activity (function$_1$) of the mammalian heart (question 1). This knowledge is typically conveyed by means of diagrams and tables. I restrict myself to a few typical characteristics of the structure of mammalian hearts. The heart is a hollow muscle containing two separate cavities (a left one and a right one), each consisting of two compartments: an atrium and a ventricle. The left channel contains oxygenated blood, which flows from the lungs to the organs; the right channel contains deoxygenated blood, which flows from the organs to the lungs. Inside each channel the blood moves from the atrium to the ventricle. The atrium and the ventricle are separated by valves. There are also valves at the end of the ventricles. The left and the right ventricle differ markedly in their structural characteristics. The cavity of the right ventricle is narrow space enclosed between two large surfaces; its wall is of moderate thickness. The cavity of the left ventricle has a cylindrical shape with a conical end, it is enclosed in a heavy layer of muscular tissue.

The causal role of the heart in the maintenance of the organism (question 2 – how is the heart used?) is well known to philosophers: the heart contributes to the organism’s capacity to circu-

---

[8] Note that functional morphology if it is defined in this broad sense is not a subdiscipline of morphology if morphology is defined as the study of form.

late blood by pumping the blood around. Note that this answer is part of a capacity explanation of a capacity of a system of which the heart is a part, namely the capacity of the circulatory system (consisting of heart, blood and blood-vessels) to circulate blood.

Subsequently it is asked how the heart works: how is it able to pump blood, how is the flow directed and how is the heart's activity regulated (question 3)? Note that these questions come down to the question 'how is the heart able to perform the causal role it has?' Hearts are, in essence, two-stroke pumps with a filling phase and an emptying phase. The blood is pushed forward by contracting the muscular tissue in the walls of the chambers. The right ventricle works like a bellows (the walls are pulled together and a moderate force is generated which works over a relative large distance), the left one like a pressure pump (the chamber is fiercely constricted). The valves assure unidirectional flow of the blood within the heart and work passively. Note that this capacity explanation attributes causal roles to the parts of the heart. For example it attributes the causal role to generate force to the muscular walls and the causal role to prevent blood from flowing in the wrong direction to the valves.

The answer to the type (4a) question concerning the utility of having (rather than lacking) the ability to pump blood seems trivial: if the circulatory system did not have a pump it would not be able to circulate blood.[9] Note that this answer refers to the causal role of the encompassing system.

Numerous type (4b) questions can be asked about the way in which the pumping role is implemented. For example, 'why is the causal role to pump blood implemented by means of special organ rather than by contracting the arteries?', 'why are there two pumps (a left one and a right one)?', 'why are the two pumps implemented in one organ?'. Other type (4b) questions concern the specific characteristics of the pump. Why does each pump consist of two parts (an atrium and a ventricle)? How to explain the structural and functional differences between the left and the right ventricle? Why does it beat with the frequency with which it beats (this frequency varies between species)? Why does it have the volume it has? Why is it important that the blood flows in only one direction? And so on ....

Type (4b) questions are answered by means of a design explanation of the character of the item or behaviour in question. Consider, for example, the answer to the question why birds and mammals have a double circulation (the blood passes the heart twice) rather than a single circulation (in which the blood would pass the heart only once). When the blood passes the capillaries of the lungs, the blood pressure drops to a level too low to get the blood through the capillaries of the body. A single circulatory system in which the heart pumps the blood directly from

---

[9] Because the answer is obvious, the question is often omitted in text books. However, the question can be made more interesting by introducing more interesting alternatives for pumping (more interesting than simply lacking a pump): why for instance is the blood circulated by means of a pump rather than by means of cilia?

the lungs to the organs would need a pump that generates much more pressure than the hearts of birds and mammals do. Such a pump would create several problems. One problem is that the capillaries of the lungs would be blown up because of the high pressure of the blood they contain. Other problems concern the design of such a pump. A double circulatory system in which the blood is routed back to the heart after it has passed the capillaries of the lungs solves the problem created by the drop in blood pressure when passing the lungs without creating the problems a much stronger pump would do. This is a typical design explanation pointing out that given the structure of the lungs the role performed by the system of heart and blood vessels is performed better if that system contains two pumps than if it contains one.

Next, consider the question why both pumps have two compartments. This is because otherwise the blood would flow not fast enough (the two chamber structure reduces the time needed to fill the pump because one part is filled while the other pumps).

Especially interesting is the design explanation of the differences between the structural and functional characteristics of the right and the left ventricle. The right ventricle pumps the blood through the pulmonary circuit (i.e. via the pulmonary artery to the capillaries of the lungs and via the pulmonary veins back to the left atrium). The left ventricle pumps the blood through the systemic circuit (i.e. via the aorta to the capillaries of the organs and back to the right atrium via the veins). The different characteristics reflect the different demands posed upon the pump by these different circuits. The pressure gradient in the systemic circuit is many times that of the pulmonary circuit. This means that the pump of the systemic circuit (the left ventricle) must generate much more pressure than the pump of the pulmonary circuit (the right ventricle). However, as the two circuits are connected in series the pulmonary circuit must transport the same volume of blood per unit of time as the systemic circuit. In order to transport the same amount of blood in the same time with less force the blood must be moved over a greater area in that time. This explains why the left ventricle works as a pressure pump and of the right one as a volume pump. It also explains the form of the cavities and the walls (a volume pump needs a larger surface to volume ratio than a pressure pump and in order to generate high pressures one needs a thick layer of muscle).

Type (5) questions concern the development of the heart in the course of the ontogeny. Heart development starts early in the ontogeny when the embryo is not yet separated from the yolk sac. The heart starts as a paired organ in the mesodermal germ layer. At each side of the body a longitudinal tube develops. As the embryo separates from the yolk sac in the pharyngeal region these two tubes fuse into each other to form a single cavity, in front of and behind this region the tubes remain separate. At the same time the tube elongates and twists resulting in four chambers in series: sinus venosus, atrium, ventricle and conus arteriosus (in the direction in which the blood flows). This resembles the structure of the adult heart in typical fish (except for the absence of valves). Further development includes: subdivision of both the atrium and the ventricle in two parts, absorption of the sinus venosus in the right atrium, separation of the pulmonary veins from the sinus venosus so that they open in the left atrium, and subdivision of the conus arteriosus into the pulmonary trunk originating from the right ventricle and the aorta originating from the left ventricle. The heart is the first organ which starts to perform its causal role. Beating begins even before the two tubes are fused. The heart's development is dependent on its own activity. If there is no blood flow, development stops shortly after the fusion of the tubes. The size of the heart causally depends on the volume of blood passing through it.

Type (6) questions concern the evolution of the heart. The vascular system of the ancestral vertebrates is assumed to resemble that of amphioxus (a non-vertebrate chordate).[10] This system is sketched in fig. 2.1.[11]

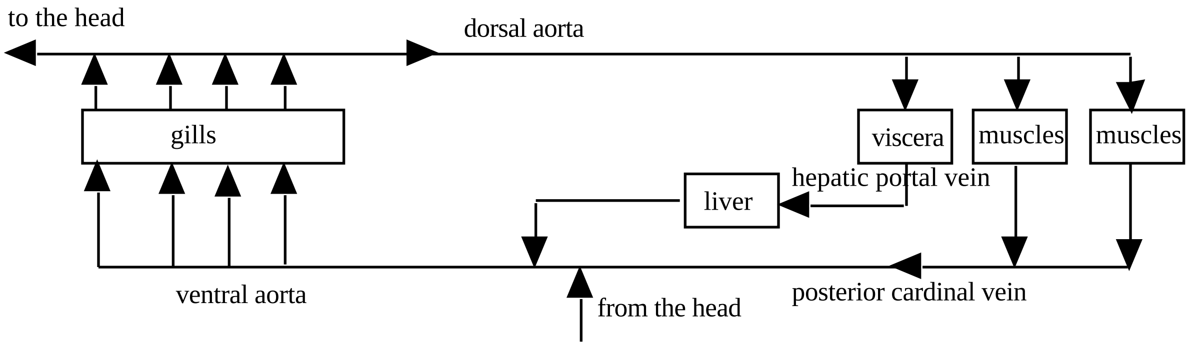

Fig 2.1: blood circulation in amphioxus

Amphioxus does not have a separate heart; its blood is propelled by peristaltic contraction of some large blood vessels in the pharyngeal region, the main one is the ventral aorta. About fifty branchial vessels branch off from the ventral aorta, pass through the gills[12] and reunite into a large longitudinal blood-vessel the dorsal aorta. The branchial vessels are enlarged at their base and these enlargements contract, thus contributing to the propulsion of the blood. From the dorsal aorta blood is routed to the vessels of the segmental muscles and to the vessels of the intestine. Blood from the segmental muscles returns via the cardinal veins to the ventral aorta. Blood from the intestine is routed back via the liver to the ventral aorta. It is supposed that the vertebrate heart developed somewhere in the ventral aorta of an ancestor which an amphioxus like vascular system, resulting in the fish-like circulation outlined in fig 2.2.

---

[10] The phylum Chordata consists of three subunits with the rank of subphylum: the Tunicata, the Cephalochordata (of which amphioxus is a species) and the Vertebrata (of which the mammals are a class).

[11] Note that the figures are highly schematized. Important things which are not represented in these schemes include the structure of the heart, the blood vessels to and from the head and the renal portal system.

[12] The gills in amphioxus and, presumably, in vertebrate ancestors have a nutritional rather than a respiratory role. Respiration in amphioxus occurs through the skin.

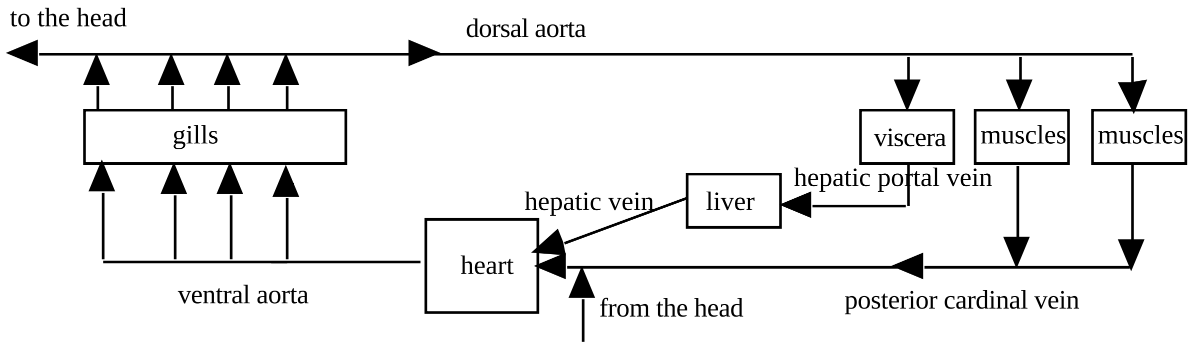

Fig 2.2: blood flow in a fish-like ancestor with a single circulation

The double circulatory system of birds and mammals evolved from such a single circulation in fish-like ancestors. The evolutionary explanation of this transition (see, for example, Johansen 1977: 374-377) points out that birds and mammals evolved from ancestors that changed their respiration from the gills to the lungs. The evolution of a double circulation in modern birds and mammals from a supposed single circulation in their fish-like ancestors started some 350 million years ago. The climate conditions of the time brought about very low oxygen concentrations in large tropical fresh-water basins. As a result many species of fresh-water fish developed a special organ for air breathing in addition to their gills.[13] These species relied on their lungs for oxygen absorption, but retained their gills for the elimination of oxygen and osmoregulation. Such a bimodal (lungs & gills) mode of respiration favoured rearrangements in the vascular system that increased the effectiveness of the new organ for gas exchange. The heart of this ancestral organism probably consisted of four chambers in series: sinus venosus, atrium, ventricle and conus arteriosus (this is the same structure as that in typical fish). The arrangement of the main blood vessels of the ancestral organism is outlined in fig. 2.2. The blood is supposed to flow from the heart via the ventral aorta to the gills and then to the dorsal aorta. The dorsal aorta is a large blood vessel which runs over the entire length of the body. From the dorsal aorta many main arteries branch off which transport the blood to the viscera (intestine) and the muscles. Blood from the viscera returns to the sinus venosus of the heart via the liver and the hepatic vein, blood from the muscles returns to the sinus venosus via the cardinal vein.

The lung evolved as a diverticulum of the foregut. This resulted in the vascular arrangement sketched in fig. 2.3. Blood is routed from the heart to the gills and then to the organs, among which the lung. Blood which is oxygenated in the lungs flows via the hepatic vein to the sinus venosus of the heart.

---

[13] Modern teleost fish too are supposedly derived from an ancestor having a lung. This lung developed into a swim bladder.

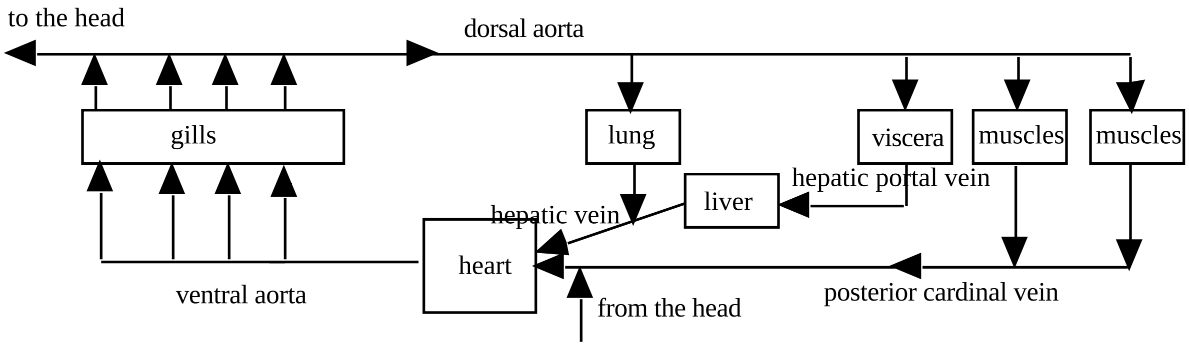

Fig 2.3: development of the lungs

Note that already at this stage oxygenated blood does not flow directly from the main respiratory organ (the lung) to the organs that need the oxygen, as is the case in a single circulation. Instead it flows from the lungs back to the heart and then to the other organs (this is a consequence of the fact that the lung evolved as a an organ in an organism in which the blood flowed in the manner sketched in fig. 2.2). As a result oxygen rich blood from the lungs mixes with oxygen poor blood from the organs. This arrangement favours structural changes in which oxygenated and deoxygenated blood are kept separate. An arrangement in which oxygenated and deoxygenated blood are separated is more efficient because in this situation the concentration gradient in the lungs as well as in the oxygen consuming organs is greater, which speeds up the diffusion process at these sites. The first step in the direction of more efficient air breathing was the development of a separate blood-vessel (the so-called ‘pulmonary vein’), routing blood from the aerial gas exchange organ (the ‘lung’) directly back to the atrium of the heart (fig. 2.4).

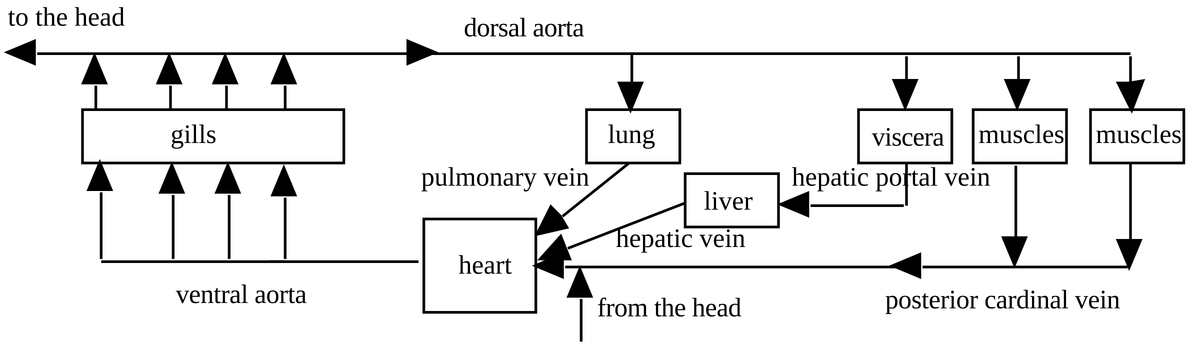

Fig 2.3: emergence of a pulmonary vein

Thus, the pulmonary vein emerged as a means to minimize the mixing of oxygenated and deoxygenated blood in an aquatic ancestor with a bimodal mode of respiration. In air breathing fish the surface area of the gills became reduced as air breathing became more important. The reduction of the gills is correlated with changes in the aortic arches. The first change was the development of a pulmonary artery (fig. 2.5).

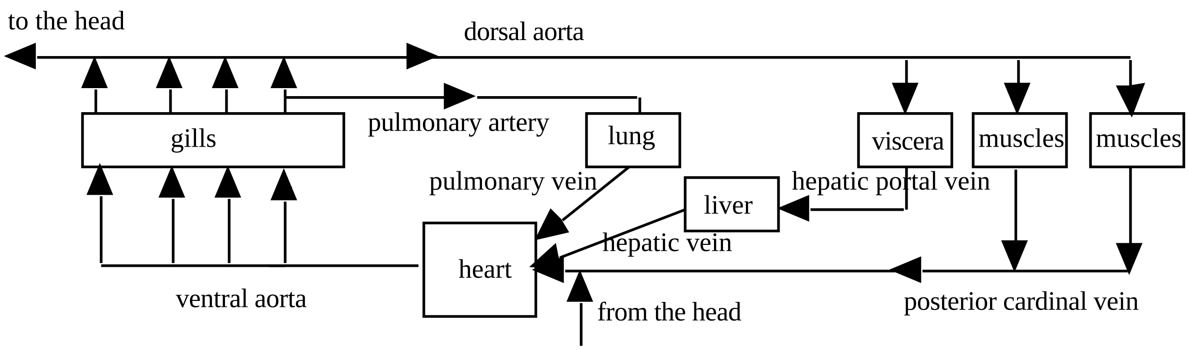

Fig 2.5: emergence of pulmonary artery

The development of specialized blood vessels routing the blood from the heart to the lung and back from the lung to the heart was accompanied by structural changes in the heart. In the stage sketched in fig 2.3 the pulmonary vein (containing oxygen rich blood from the lungs) enters the heart in the left part of the atrium, whereas the hepatic and cardinal veins (containing oxygen poor blood from the organs) enter the heart in the sinus venosus. The left and the right part of the atrium are partially separated by means of a septum. The left part receives oxygenated blood from the pulmonary vein, the right part receives deoxygenated blood from the sinus venosus. In the ventricle two partial septa develop (from opposite sites) which separate the left and the right part of the ventricle. The conus arteriosus modifies such that deoxygenated blood from the right part of the ventricle is shuttled to the branch of the aorta from which the pulmonary artery departs and oxygen rich blood from the left part of the ventricle is shuttled to the other branches. In the subsequent history the sinus venosus is gradually incorporated into the right atrium, the septum in the atrium becomes complete, the two ventricular septa fuse, and the conus arteriosus reduces. This results in the four chambered heart found in birds and mammals. In the circulatory system of these organisms oxygen rich blood and oxygen poor blood are completely separated (fig. 2.6).

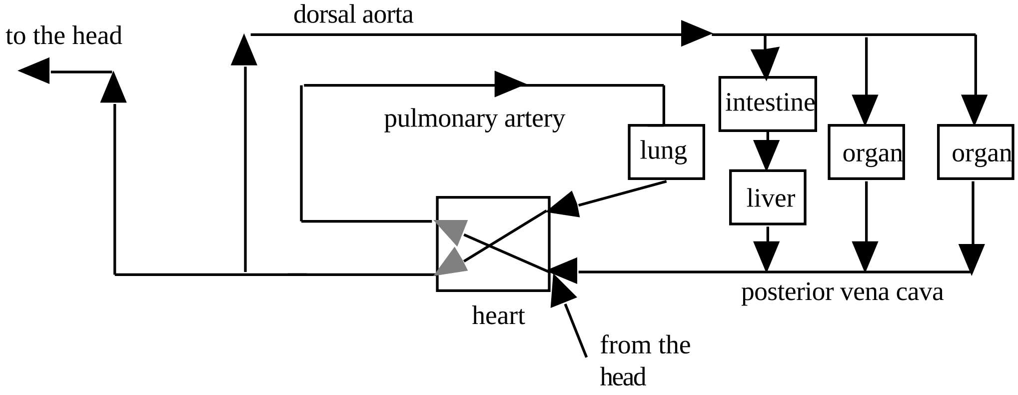

Fig. 2.6: double circulation in birds and mammals

Ethology

Ethology is often broadly defined as “the biological study of animal behaviour”. It’s main concerns are:

1) the description of behavioural patterns,
2) the causal role of those behavioural patterns in the maintenance of the organism,
3) the causation of occurrences of behavioural patterns and the mechanisms by which behaviour is produced,
4a) the survival value of performing certain causal roles,
4b) the survival value of the different forms of behaviour,
5) the ontogenetic development of behaviour,
6) the evolution of behaviour.

I shall discuss these questions with an emphasize on the singing behaviour of birds.

The empirical foundation of ethology consists of accurate descriptions of the behaviour of the organisms under study (the so-called ‘ethograms’). These descriptions may concern both the structural aspects of behaviour (such as the order in which certain activities are performed) as the functional (function$_1$) aspects (how fast does a leopard run? how many fish does a gull catch within an hour?). Beginning students are warned not to make hypotheses about the causal role of the behaviour when they describe behaviour. Singing is a typical example of a behaviour that is characterized in terms of its structure. Bird sounds are traditionally divided into calls and songs. One of the main distinctions between these two is that songs are much more complex in structure than calls.

The second kind of problems (type (2) questions) concerns the causal role of the described behavioural patterns in the maintenance of the organism. After the behaviour is described one may hypothesize about the causal role of that behaviour in some larger context. For example: the horse’s habit to pile up their faeces and the fox’s habit to urinate at certain places have a causal role in keeping a territory; the gull’s habit to catch fishes has a causal role in feeding and so on. According to Catchpole (1979: 31-37) the two main causal roles of singing in birds are attracting females and repelling rivals (other possible causal roles are synchronizing the reproductive behaviour of a pair and maintaining the strength of the pair-bond). In saying that the singing attracts females one describes how the singing contributes to the capacity to reproduce, in saying that the singing repels rivals one describes how the singing contributes to the capacity to maintain a territory. The song acts as a first line of defence. If singing does not work (e.g. because the birds are muted by an experimenter) invaders are expelled by visual display and actual attack. There appears to be an interesting difference between the causal roles of singing in reed warblers and in sedge warblers: in reed warblers singing has both a sexual and a territorial role, whereas in sedge warblers it has only a sexual role.

Type (3) questions are concerned with the generation of behaviour: the external causes (stimuli), the internal causes (motivation), the mechanisms that link sensory information to behaviour and the way in which the behaviour is organized. The explanations is this areas of research are usually called ‘causal explanations’ by ethologists. An example is the explanation of why a certain male chaffinch starts to sing in the spring by pointing to the increasing length of the day. These explanations are of the same kind as ‘physiological explanations’ in functional morphology, and I will use that latter term to avoid confusion with other uses of the term ‘causal explanation’.

The next two kinds of problems (type 4a and type 4b questions) concern the survival value of behaviour. Explanations that appeal to survival value are usually called ‘functional explanations’ or ‘ecological explanations’ by ethologists. I shall use the term ‘design explanations’. Design explanations in ethology explain why a certain behaviour is performed in the way it is performed in terms of the utility of performing this behaviour to the organism that perform it. This utility is related to the other traits of the organism and the state of the environment in which they live.

Type (4a) questions concern the utility of performing the roles performed by the behaviour under study. After having concluded that singing has a causal role in reproduction and in maintaining a territory one continue and ask ‘why is it useful to reproduce?’ and ‘why is it useful to defend a territory (rather than occupy an undefended living space)?’. The answer to the first questions is obvious and trivial.[14] The answer to the second questions is an important subject of research.

Type (4b) questions concern the way in which those roles are implemented. An example is Catchpole’s (1979: 31-37) explanation of why in reed warblers singing has both a sexual and a territorial role, whereas in sedge warblers it has only a sexual role. This difference is explained by means of a design explanation by pointing to the different environmental conditions:

> Reed warblers inhabit dense, impenetrable reed beds where vision is poor, and have small territories which need to be constantly defended by vocal means. In contrast, sedge warblers are more scattered throughout open, terrestrial habitats, where vision is good and less premium placed upon focal defense (Catchpole 1979: 33)

In other words: given the environmental conditions in which reed warblers live they need to defend their territory by singing, whereas given the environmental conditions in which sedge warblers live, they can do without.

---

[14] However, the answer to the question why reproduction in e.g. birds takes the form of sexual reproduction (a type (4b) question) is not obvious and far from trivial.

Type (5) questions ask for an explanation of the behaviour under study as the result of the interaction between genetic and environmental influences. It was Thorpe who started the study of song development in the late fifties (Thorpe 1958). If a young chaffinch is reared without ever hearing a singing chaffinch it will never be able to sing the full song. This shows that the young chaffinch must *learn* the song from its father. Thorpe shows that the juvenile learns the song in the first few months of its life, long before it is able to produce that song itself. The period during which the young is able to learn the song is called the ‘critical period’. The young chaffinch does not learn any song it hears in its youth. It picks up only the songs of the chaffinch and songs that are very much like those of the chaffinch, such as that of the willow warbler and the tree pipit. Hence, there must be an *innate* component that enables the juvenile to recognize the song of its father as one that it should learn. Marler (1975: 24-29) explains this pattern of development, that is manifested by the chaffinch and many other birds, by hypothesizing the existence of an inherited, but modifiable, auditorial template. During the critical period this template is improved in view of the adult model. The template itself determines which songs are accepted as a model. The improved template specifies which dialect the bird has learned. When the young starts to sing it first produces the so-called ‘subsong’, which differs from the full song in a number of ways. It is more quiet, has a much wider range of frequencies and has little or no structure. The vocal output is gradually matched to the dialect specified by the improved template until the bird manifests its full song.

Not much is known of the evolutionary history of bird songs (type 6 questions). An intriguing question in this area is the question why songs have become so complex during evolution. Bird sounds are traditionally divided into calls and songs. One of the main distinctions between these two is that songs are much more complex in structure than calls. Yet, simple calls are quite capable of carrying the information that is conveyed via songs. Indeed, only songbirds have evolved complex songs and all other birds manage to keep a territory and find a mate without singing. Several hypotheses have been suggested to explain the evolution of complex songs, the main ones are based on the mechanism of sexual selection. Suppose that due to some cause or another female birds prefer vocalizations that are more complex than average. As a result, mutant males with more elaborate songs will produce more offspring and the complexity of the song will gradually increase. The female preference might be just a quirk of brain chemistry (this is the so-called “runaway theory of sexual selection”). Another hypothesis is that the preference of females results from natural selection. On this hypothesis male birds that are able to sing elaborately in the morning show that they have enough energy resources left after a night of fasting. Hence, female birds who choose a male that sings elaborately choose a male that is fitter in terms of his ability to look after offspring and will, therefore, produce more offspring (Hutchinson, McNamara & Cuthill 1993).

Conclusion

I have argued that functional biologists ask seven types of questions in relation to the item or behaviour in which they are interested. These questions concern: (1) the character of the item or behaviour under study, (2) their causal roles, (3) the causes and underlying mechanisms resulting in the realization of those causal roles, (4a) the survival value of performing those, (4b) the survival value of the character of the item or behaviour under study, (5) the way in which the item or behaviour under study develops in the course of the ontogeny, (6) the way in which the item or behaviour under study developed in the course of evolution. These questions may be phrased as follows:

(1) what is the character of the item or behaviour under study?
(2) what are its causal roles?
(3) how does it perform those causal roles?
(4a) why is it useful to the organism that it has an item that performs those causal roles?
(4b) why is it useful to the organism that the item or behaviour in question has the character it has?
(5) how did that item or behaviour develop in the course of the ontogeny?
(6) how did that item or behaviour develop in the course of the evolutionary history?

Note, that the type of question is determined by the kind of problem it addresses, not by the words used to phrase that question. In the overview above I have phrased these questions in such way that they best represent the issues at stake. Of course biologists may use and do (often) use other words to phrase these questions. Biologists typically use the phrase ‘what is the function of [item] ....?’ (for example ‘what is the function of the thymus?’) to phrase a type (2) question. The phrase ‘why do .... organisms have [item] ...?’ (for example ‘why do vertebrates have a circulatory system?’) is often used to phrase a type (4a) question, and the phrase ‘why does [item / behaviour] .... of .... organisms have [character] ....?’ (for example ‘why do snakes have a forked tongue?’) is often used to phrase a type (4b) question. The why-questions as they are phrased by biologists are notoriously ambiguous. For example, biologists use the phrase ‘why do .... organisms have [trait] .....?’ not only to ask for design explanations but also to ask for physiological or evolutionary explanations.

In order to avoid confusion it would be best to phrase the questions in the manner in which I have phrased them in the overview above. However, I will sometimes use the ‘why do .... organisms have [trait] .....?’ mode in order to connect the things I say to things others have said. In this case I will use subscripts to distinguish the several types of why-questions. I distinguish:

1) questions that ask for causes at the level of an individual organism (why$_1$-questions)

2) questions that ask for the utility of a certain trait to the organisms that have it (why$_2$-questions)

3) questions that ask for evolutionary causes (why$_3$-questions)

Why$_1$-questions ask for physiological and/or developmental explanations, why$_2$-questions ask for design explanations and why$_3$-questions ask for selection explanations.

2.3.3 The role of the different kinds of function statements

Descriptions of (potential) activity (function$_1$)

Function$_1$ attributions (description of functional characteristics) have the same role in biological research as descriptions of structural characteristics.

The structural (form) and functional (activity) characteristics of an organism, its parts and behaviours are to be explained by means of four kinds of explanations:

- physiological explanations (explain how a certain activity or behaviour is brought about in certain organisms, that is they answer type (3) questions);
- design explanations (explain why it is useful to the organism that the item or behaviour in question has the functional and structural characteristics it has—answer type (4a) and (4b) questions);
- developmental explanations (explain how the functional and structural characteristics of an item or behaviour are brought about in the ontogeny—answer type (5) questions);
- evolutionary explanations (explain how the functional and structural characteristics of an item or behaviour are brought about in the evolutionary history—answer type (6) questions).

These four ways of explaining an item or behaviour are complementary. In most cases a complete explanation of an item or behaviour of an organism would include all four kinds of explanations. There are, however, exceptions. For example, if an item does not have a causal role (think of the human veriform appendix) or if a certain character does not have survival value (think of the colour of vertebrate bones) there is no need for a design explanation.

Furthermore, functional and structural characteristics are used to explain the utility of other functional and structural characteristics (of the same organism) in a design explanation of those latter functional and structural characteristics.

Attributions of causal roles (function$_2$)

It will be clear from the discussion in the preceding sections that the notion of function as causal role (function$_2$) is central to functional biology. It is the handle by means of which functional biologists get a grip on their subject matter:

(i) Attributions of causal role tell us which tasks a certain item or behaviour has. They answer a type (2) question (how is that item used?). The tasks of an item or behaviour are to be explained by means of a capacity explanation (in answer to a type (3) question – how is that item or behaviour able to perform those tasks?) and a design explanation (in answer to a type (4a) question – why is it useful to the organism to have an item or behaviour which performs those tasks?).

(ii) Attributions of causal role serve to explain how an item or behaviour is able to perform the tasks it has by means of capacity explanations (in answer to a type (3) question). Capacity explanation work by attributing causal roles to the components of the item or behaviour the capacity of which is to be explained.

(iii) Attributions of causal role serve to asses survival value as part of design explanations of the character or presence of an item or behaviour (in answer to type (4b) questions). Such design explanations explain why an organism is built the way it is, why it works the way it works, or why it behaves the way it does, by appeal to the utility of those items and behaviours to the organisms that have it. They typically proceed by showing (1) that the item or behaviour in question has some causal role in the maintenance of the organism; (2) that given the way in which the organism is built, the way it behaves and the state of the environment in which it lives, that role is better performed by an item or behaviour with the character in question than by other conceivable forms.

(iv) Attributions of causal role serve to explain the evolution of an item or behaviour by showing that that item or behaviour evolved because some variant appeared in which that item or behaviour performed its causal role better (in answer to type (6) questions).

Function as survival value –design explanation (function$_3$)

Claims about survival value are an essential part of design explanations. Design explanations answer type (4a) and type (4b) questions. They explain why the character or presence of the item or behaviour under study is useful to the organisms that have it. Design explanation is one of the four complementary ways in which biologists explain items and behaviours. Design explanations of the character or presence of an item or behaviour of a certain organism compare real organisms with hypothetical organism in which that item or behaviour has another structure, or in which that item or behaviour is absent or replaced by another item or behaviour.

Function as selected effect–selection explanation (function$_4$)

Functional biologists seldom or never use the term ‘function’ in the sense of function as selected effect.[15] I have nevertheless included this notion in my list because it is important to

---

[15] Please note that I argue for this claim in chapter 7.

distinguish this notion clearly from the other notions of function. This is important because (1) the notion of function as selected effect seems to be the notion of function with which philosophers who favour an etiological account are concerned, and (2) there is at least one evolutionary biologist (Williams) who maintains that function should be defined in historical terms.

Analyses of biological function in historical terms have added nothing but confusion: such analyses suggest that they are concerned with the clarification of an existing use of the term ‘function’, whereas in fact they give an existing term a meaning it never had before. In other words such analyses introduce a new meaning for a term which was already highly ambiguous. This is a bad way of doing philosophy.

Of course, when I say that analyses of function in historical terms add nothing but confusion I do not mean to say that biologists do not or should not study function in an evolutionary context. On the contrary, as I will show, appeals to functions are an important part of explanations which explain why and how a certain item evolved. Furthermore, change of function in the course of evolution is an important mechanism of evolutionary change. My point is rather that when biologists use the term ‘function’ in the context of an evolutionary study they use that term in a non-historical sense. They aim to refer to the causal role (function$_2$) of an item or behaviour or to the survival value of a certain trait but not to function in some fancy historical sense. Definitions of function in historical terms obscure this fact.

2.4 Summary and conclusion

In this chapter I sketched the practice of functional biology. I showed that the term ‘function’ is used in at least four different ways. It may refer to (1) the activity of an item or behaviour (function$_1$), (2) the causal role of an item or behaviour (function$_2$), (3) the survival value of having a certain character or performing a certain role (function$_3$), (4) the effects for which a certain trait was selected in the past (function$_4$). To understand how these notions are used in functional biology, I discussed the aims of research in functional biology.

I showed that functional biologists try to answer seven different types of questions about the form (structure), function$_1$ (activity characteristics) and behaviour of organisms and their parts. These questions concern: (1) the character of those items and behaviours, (2) their causal roles in the maintenance of the organism; (3) the mechanisms by means of which these causal roles are performed, (4a) the survival value of performing these causal roles, (4b) the survival value of having a certain character, (5) the development of those items and behaviours in the course of the ontogeny, (6) the evolutionary history of those items and behaviours.

Let me review these questions with respect to their relevance to explanation (after all, my main concern is the role of appeal to function in explanation in functional biology). Functional biologists aim to explain the form, function$_1$ and behaviour of organisms and their parts. They do so in four different and complementary ways: by means of physiological explanations, design explanations, developmental explanations and evolutionary explanations. Physiological explanations answer questions of type (3). They explain how a certain item or behaviour works, that is how it performs the tasks it has. A special kind of physiological explanation is capacity explanation. A capacity explanation is a physiological explanation that explains the capacity of an item or behaviour to perform a complex task by attributing to the parts of that item or behaviour the ability to perform a series of less complex tasks that add up to the capacity to be explained. Design explanations answer questions of type (4a) and (4b). They explain why it is useful that a certain item or behaviour has a certain character or why it is useful that a certain role is performed by relating this trait to the other traits of the organisms that have that trait and to the state of the environment in which they live. Developmental explanations answer questions of type (5). They explain how a certain item or behaviour develops in the course of the individual's history. Evolutionary explanations answer questions of type (6). They explain how a certain item or behaviour evolved in the course of the history of the lineage. A special kind of evolutionary explanation is evolutionary selection explanation. An evolutionary selection explanation is an evolutionary explanation that focuses on natural selection.

In order to explain the form, function$_1$ and behaviour of a certain type of organism one needs to know what the form, function$_1$ and behaviour are. One reason why type (1) questions are relevant to explanation is that the phenomena described in answer to type (1) questions are to be explained by means of physiological, design, developmental and evolutionary explanations. Of course, form, function$_1$ and behaviour are not only important in posing questions about mechanisms, design, development and evolution, but also in answering them. This is another reason why type (1) questions are important to explanations.

Questions of type (2) are relevant to explanation because attributions of causal roles have a key role in the strategies of explanation in functional biology. The causal role of an item or behaviour is important in at least three kinds of explanation: physiological explanation, design explanation and evolutionary explanation. Attributions of causal role give biologists a handle that enables them to get a grip on these explanations and to connect them. Attributions of causal roles specify the tasks a certain item or behaviour fulfils. These task define what is to be explained by means of physiological explanations (how does an organisms perform a certain task?). Design explanations are concerned with the question why these tasks are performed and why they are performed in the way they are performed. Evolutionary selection explanation explain character change by pointing out that somewhere in the past a new variant emerged which performed a certain task better than the old variant (in the conditions in which those organisms lived).

Chapter 3: Biologists about function

3.1 Introduction

Biologists have written surprisingly little on the notion of ‘function’. Although they often investigate functions they seldom delve into the question what it is to be a function (at least not in writing). The two main discussions of ‘function’ in morphology are still those of Bock and Von Wahlert (1965) and of Dullemeijer (1974). Other important discussion can be found in Zweers (1979) The main discussion of the notion of function in ethology is that of Tinbergen (1963).

Table 3.1: how different morphologists use certain words related to form and function.

	Item	Form	Activity	Role	Survival value
Bock & Von Wahlert (1965), Bock (1980)	feature	form	function	biological role	
Dullemeijer (1974)	element	form	(potential) action, activity	meaning, purpose, role, significance	
Zweers (1979)	element / (sub)system	structure	action	role (function, meaning)	
Lauder (1986, 1990)	component	structure	function (morphology)	function (ethology)	

3.2 Bock and Von Wahlert (1965)

The single most important attempt to clear up the notion of ‘function’ in morphology is made by Walter Bock and Gerd von Wahlert in their “Adaptation and the form-function complex” (1965). In this paper they contend that the term ‘function’ has two different meanings, namely ‘what it does’ and ‘what it is used for’. They propose to restrict the meaning of ‘function’ to the first meaning and coin the expression ‘biological role’ for the second. Bock & Von Wahlert are impressed by the linguistic turn in philosophy and they propose to make the distinction between function (what an item does) and biological role (how it is used) in terms of the kind of predicates involved in statements describing an organism. In their view, the difference between what an item does (function) and how it is used (biological role) is a difference of whether or not the predicates used to describe the function or role refer to the environment or not. (Bock & Von Wahlert do not explain what it is for a predicate to refer to the environment or not, but I take it that they mean that the definition of that predicate refers to the environment.) Biological roles are, by definition, the activities of organism described in terms of predicates that refer to the environment (p. 278). Predicates that do not refer to the environment describe forms or functions. ‘Form’ refers to “the class of predicates of material composition and the arrangement, shape or appearance of these materials, provided that these predicates do not mention any reference to the normal environment of the organism” (p. 272). ‘Function’ refers to “that class of predicates which include all physical and chemical properties arising from its form [...] including all properties arising from increasing levels of organization, provided that these predicates do not mention any reference to the environment of the organism” (p. 274). Bock & Von Wahlert emphasize that a certain function in one animal may have many different roles and that the same function may have different roles in animals of different species:

> For example, the legs of a rabbit have the function of locomotion—either walking, hopping, or running—but the biological roles of this [activity][1] may be to escape from a predator, to move toward a source of food, to move to a favorable habitat, to move about in search of a mate, and so forth. The legs of a fox also have the function of locomotion, although the details of the form and the function of the fox leg differ greatly from the rabbit leg. Yet some of the biological roles are quite different in the fox and in the rabbit. One role of the leg in the fox would be to catch its prey when it is chasing the rabbit in which the role would be to escape from its predator (Bock & von Wahlert 1965: 279).

According to Bock & Von Wahlert the distinction between function (what a certain item does) and biological role (how it is used) is important for two reasons. First, they want to emphasize that it is not sufficient to study an organism’s capacities in order to know how those capacities are used by the organism. As an example they recall Bock’s mistake in inferring the use of the mucus gland of the gray jay (*Perisoreus canadensis*):

> Bock (1961) suggested that the biological role of the large mucus glands of the gray jays was to coat the tongue with a glue-like material which would allow the bird to obtain food from crevices in the bark of trees and thus is similar to the biological role of the large mucus glands in some woodpeckers. Dow (1965), however, has shown, by observations of these birds in captivity and in the wild, that the mucus serves as a glue to cement food particles together into a food bolus which is then stuck to branches of trees. These food boli are a device to store food during the winter; the stored boli are found and eaten

---

[1] Bock & Von Wahlert use the term 'faculty', which is a combination of a form and a function$_1$.

> during stormy weather and other periods during which these jays cannot find food (Bock & von Wahlert 1965: 278/9).

Bock & Von Wahlert’s second reason is related to their first: they seek to emphasize the importance of field studies to morphology.

> The function of [an item][2] may be studied and described independently of the natural environment of the organism as is done in most studies of functional anatomy (Bock & von Wahlert 1965: 274)

> Essential to the description of a biological role is the observation of the organism living naturally in its environment. The descriptive adjective “biological” stresses this fundamental property of the biological role. A biological role cannot be determined by observations made in the laboratory or under artificial conditions. (Bock & von Wahlert 1965: 278)

Functions in their sense are the things that can be studied in the laboratory, biological roles should be studied in the wild. Bock has repeated this concern several times (1980, 1990). He also urges the need for developing and funding the morphological study of the relation between organisms and environment (ecomorphology) on the ground that with the extensive destruction of natural environments this study may soon become impossible.

In emphasizing the distinction between “what an item does” and “what an item is used for” Bock & Von Wahlert make an important point that should be taken into account in any account of function. In my account this distinction appears as the distinction between on the one hand an item’s activities and capacities (function$_1$) and on the other hand its causal roles (function$_2$). However, the way in which Bock & Von Wahlert seek to make this distinction has several problems. As I said, they seek to make this distinction in terms of whether or not the predicates involved refer to the organism’s environment. This does not work for reasons that will become clear shortly. Moreover, their association of this distinctions with the distinction between what can be studied in the laboratory and what should be studied in the field does not work either.

A first problem is that in stating how an item is used one often does not refer to the organism’s environment. In section 2.1.1 I discussed the circulation of the blood (example 2.1). This is an activity (something an item does) that is used by the organism for, among other things, the transport of oxygen, nutrients, wastes and heat. In saying that the circulatory system is used to transport oxygen, nutrients, wastes and heat one does not refer to the environment of the organism. Bock & Von Wahlert themselves use the term ‘role’ several times in this sense:

> Bony skeletal elements possess great strength against compression, tensile and shearing stresses which allow their *role* as support of the body, points of attachment for muscles and mechanical protection of vital organs (Bock & von Wahlert 1965: 275, italics mine).

---

[2] Bock and Von Wahlert use the term 'feature'

> Collagenous fibered tendons and ligaments possess great strength against tensile stresses and are highly non-compliant (not stretching), but they have no strength against compression or shearing forces which allows their *roles* as the intermediate structure between muscles and bones and the ties between individual bones (Bock & von Wahlert 1965: 275, italics mine).

This way of talking suggests a tripartion along the lines I have set out in chapter 2.

A second problem is that many predicates that tell us what an item does do refer to the environment of the organism. Consider for instance Bock & Von Wahlert's contrast between the "function" of locomotion and the "biological role" of escaping from a predator (see the first quote in this section). If locomotion is defined as the ability to move or as the act of moving from one place to another it clearly refers to the organism's environment. Hence, the presence or absence of reference to the organism's environment does not work to make the distinction between locomotion as a predicate that states what an item does and escaping from a predator as a predicate that states how that item is used. Other "function" predicates that clearly refer to the environment are camouflage (which may have the "biological role" of escaping from predators, as in the case of the peppered moth, or of hiding oneself to approach a potential prey as is the case with the stripes of the tiger) and insulation. Moreover, most if not all "functions" (capacities) depend on the environment. A bird's capacity to fly, for instance, depends on atmospheric pressure, drag, gravity and so on. The capacity to circulate oxygen depends on atmospheric pressure, the partial pressure of oxygen in the environment (whether it be air or water) and so on. Of course, Bock & Von Wahlert are themselves well aware that "function" depends on the environment, but they do not seem to realize that this blurs the way in which they make the distinction between "what an item does" and "how that item is used".

As I said, Bock & Von Wahlert define the distinction between "what an item does" and "how that item is used" in terms of an internal / environmental distinction. In doing so they associate the distinction between "what an item does" (its function) and "how that item is used" (its biological role) with a distinction between what can and what cannot be studied in the laboratory. Because descriptions of "functions" do not refer to the environment "functions" might be studied in the laboratory and because descriptions of "biological roles" do refer to the environment these should be studied in the wild. In emphasizing the need for field studies Bock & Von Wahlert have an important point. However (and this is the third problem of their approach) the distinction between the different kind of functions does not run along the same lines as the distinction of what can be studied in the laboratory and what should be studied in the wild. How much of the role / survival value can be studied in the lab depends on how much of the natural conditions are carried over to the lab. For a whale or a bird this will be very difficult. But what about a flatworm or a gastropod? Furthermore, some functional$_1$ data (such as the running speed of a leopard when chasing a prey) could be obtained from field studies only.

In my account the distinction between kinds of predicates is derived from a prior distinction of kinds of functions, rather than the other way round. The distinction between predicates like ‘locomotion’ and ‘camouflage’ on the one hand and predicates like ‘escaping from predators’ and ‘approaching a prey’ on the other hand, is not that the last refer to the environment and the first do not, but rather that the first refer to an observable process or state (function$_1$) and the second to a certain role in maintaining a capacity, process or state (function$_2$). By labelling a certain behaviour as ‘escaping from a predator’ one points out a role of that behaviour (for example in the survival of the organism), by labelling the same behaviour as ‘locomotion’ one describes the process but does not ascribe a role to it. One might construct role predicates from descriptive predicates by prefixing the descriptive predicates with something like ‘has a role in’, ‘takes part in’ or ‘participates in’, e.g. ‘participates in locomotion’. As a result descriptive predicates are used both in sentences that describe or label activities and capacities and in sentences that ascribe roles (in more or less detail). For example, if someone states that “rabbits have three types of locomotion: walking, hopping and running” the predicates ‘walking’, ‘hopping’ and ‘running’ are used to describe the activity of locomotion. However, the sentence “the legs have the function of locomotion” does not describe a certain activity but rather states that certain items (the legs) take part in the complex activity of locomotion. For the sake of clarity it would be better to use constructions like “the legs participate in the activity of locomotion” or “the legs have a role in locomotion”.

3.3 Dullemeijer (1974)

In his *Concepts and Approaches in Animal Morphology* Dullemeijer (1974) undertakes an investigation of the methodological and conceptual principles of functional animal morphology. This monograph is most interesting for my purposes. It shows that my account of functional explanation is close to the ideas embraced by working biologists. On the other hand my account improves on that of Dullemeijer. According to Dullemeijer the principle object of functional morphology is the relation between form and function. Form is defined as the spatial extension of an item. In the case of bone (which is Dullemeijer’s main example) the main aspects of form are: its presence or absence, its position, its size, its shape, its structure and its composition. Dullemeijer’s account of function, is less clear. According to him one should distinguish three aspects of function:

> the concept of function denotes action, biological significance and the encompassing conception of relation (Dullemeijer 1974: 49).

Action (also called “activity”) is defined as any change of form (including a change in position). Biological significance (also called ‘meaning’, ‘purpose’ and ‘role’) is the way in which an item serves the maintenance of the organism, that is its role in the animal's physiology. "The encompassing concept of relation" refers to the way in which the several items of an organism are related to each other in the organism's physiology. According to Dullemeijer:

> [An item that has a function][3] forms an essential part, a member of that existing organism, because of its relation to other parts. Function here is equivalent to the relation of one element to another (Dullemeijer 1974: 48).

This relation is constitutive of the item that has the function.

> [Items][4] must not be treated as things having properties, but as phenomena that exist consequent on relations (Dullemeijer 1974: 48).

In other words:

> function equals relation (Dullemeijer 1974: 48).

For a start, Dullemeijer considers "a statement on a relatively simple relation between form and function": "aquatic vertebrates have fins to move or to propel" (p. 52).[5] He observes that this statement is unsatisfactory:

> the simple statement on the aquatic vertebrates 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 [emphasis in original] (Dullemeijer 1974: 53).

The relation between form and function can be written symbolically as a mathematical function:

$$S = f(F)$$

where the S (of structure) stands for a form, F for a function and f for the relation of form and function (p. 54).

According to Dullemeijer there are two main approaches to determine f: "the comparative or inductive method" and "the non-comparative or deductive method" (p. 55). The comparative method correlates the form of a certain item with the function(s) it performs in the different taxa. The deductive method starts with an analysis of the requirements imposed on an item by the function(s) it performs. Next, a theoretical form (a so-called "model") is deduced that fits these requirements and this theoretical form is compared to the actual one.

The first task of functional morphology is to describe the relation between function and form. According to Dullemeijer this relation is of "an acausal character":

---

[3] Dullemeijer says at this point "a structure, an element that has a function".

[4] Dullemeijer: "elements"

[5] Note that this statement can be converted into the philosopher's standard form: "the function of fins in aquatic vertebrates is to move or to propel".

> It is a relation of correlation instead of causality, because the relation does not contain a time parameter (Dullemeijer 1974: 65).

He continues:

> After having established that such a relation does occur, the following questions arise: how is the relation to be explained in terms of underlying mechanisms or factors, how in terms of the biological role or meaning, and how has it evolved in the ontogeny and evolution (Dullemeijer 1974: 65).

In the course of the book it becomes clear that in Dullemeijer's view the relation that should be used to explain the relation of form and function (e.g. the relation between propulsion and fins) in terms of their role is that of "demand". According to Dullemeijer (p. 79) "activities are demands upon form" and it is by appeal to those demands that it becomes possible to explain the relation between form and function in terms of its role.

Dullemeijer lumps too many things together, but under the woolly language he makes some important points. Let me try to clear up his account by making a number of distinctions. Instead of speaking of "*the* relation of form and function" it is important to distinguish between several kinds of form-function relations. Firstly, there is the relation of performing a certain activity (function_1), e.g. the heart beats. This is a relation between an item (not: a form) (e.g. the heart) and an activity (beating). Secondly, there is the relation of having a certain causal role (function_2) e.g. teleost fish use their extremities[6] to propel themselves in water. This is a relation between an item (not: form) (e.g. the extremities of teleost fish) and a causal role (e.g. to propel) in individual organisms of a certain taxon. Thirdly, there is the factual correlation of the form of an item and the functions (activities and roles) it performs in different groups of organisms (e.g. all aquatic vertebrates use fins for propagating themselves). This relation is established by comparison. Fourthly, there are the causal relations that explain how the item is able to perform its functions (roles and activities) in a certain environment, e.g. how thrust is generated by beating the fins in water. Fifthly, there is the relation of the demands imposed on the form of an item by the combination of the functions (activity or role) that item performs and the environment(s) in which the organism lives. This relation can be used to explain both the form of an item (not: to explain the relation of form and function, as Dullemeijer says) in a certain individual by appeal to the function of that item and the correlation of form and function in different groups of animals (an example would be a physical explanation of why in water propagation by means of fins works better than propagation by legs). Such explanations proceed by showing that only certain forms fit the demands (requirements) imposed on it by

---

[6] I use the word "extremities" at this point because the word "fin" is ambiguous: it is used to denote both a certain item (the extremities of fishes) and the form that item takes in fishes (fins as opposed to legs). I'll return to this issue in chapter 4.

these functions. These explanations are the kind of explanations that are traditionally called ‘functional explanations’ and which I have called design explanations (i.e. the explanations ad. (3) of section 2.3).

Given these different kind of relations it is clear that one should distinguish the following activities: (1) determining which functions a certain item performs, (2) determining a correlation between form and function in different groups of animals, (3) determining which demands the functions of an item impose on the items that perform those functions, (4) explaining how a certain item performs a certain function, (5) explaining why an item is built the way it is built by appeal to the demands imposed on that item by the functions it performs, and (6) explaining a correlation between form and function in different groups of animals by appeal to the demands imposed on the items that perform those functions. Dullemeijer is himself more or less aware of these distinctions. However, he fails to state them explicitly and this confuses his account. For example, as I said a few paragraphs earlier, according to Dullemeijer there are two main approaches to determine the relation between form and function: inductive or comparative and deductive or non-comparative (p. 55). This is confused both because the comparative method and the deductive method are concerned with different kind of form-function relations (the comparative method is concerned with the factual correlation of the form of an item and the functions it performs in different groups of organisms and the deductive method is concerned with the relation of demand) and because these two relation do not exhaust the possible form-function relations (there are also the relation that a certain item performs a certain function and the relation how it performs that function). Dullemeijer is himself aware of the fact that the comparative method and the deductive method are concerned with different relations and activities where he says that

> the established relation [established by comparison] can be taken as an explanation. But one needs a deduction to explain the relation (Dullemeijer 1974: 77).

This is a confused way of saying that the comparative method establishes (but does not explain) a correlation between form and function (activity 2), and that we need the deductive method to explain this correlation (activity 3 and 5).

3.4 Zweers (1979)

According to Zweers (1979) morphologists should view their object (organisms and their parts) as systems consisting of subsystems. Morphology is basically concerned with the question “why is a system built the way it is, and why not different?” This question has three aspects. Firstly, one should ask “how is the system built” or “how does it look like”. Next, one should raise the question “how does the system work, i.e. function?”. Finally, one should ask "as a result of what evolutionary and developmental processes did the system arrive at its present state?" (p. 422). Zweers's account could be improved by making a clear distinction between "functional explanations" and answers to the question "how does it work". The first explain why a system (item) is built the way it is in terms of its roles in a larger system. The latter explain how the system (item) is able to perform a certain activity (function$_1$) in terms of an underlying mechanism. As I said in section 2.3.4 explanations of the latter kind attribute roles (functions$_2$) to the parts of the system the activity of which is explained.

Zweers contrasts "structure" (form) parameters with "action" (function$_1$) parameters:

> There are two categories of parameters which have been used in functional morphology to describe the investigated object. These categories are the structure parameters and the action parameters. The structure parameters are distance, position, size, shape, hardness, volume, mass, weight, colour, etc. The action parameters and their derivatives form a long series: movement, velocity, acceleration, vibration, force, momentum, work, power, potential and kinetic energy, friction, elasticity, viscosity, stress, strain, tensile strength, compressive strength, impact strength, heat parameters, conductivity, electric activity other electric parameters, sound, etc. (Zweers 1979: 423).

The term "role" gets its meaning in the context of systems of subsystems:

> The term "role" is primarily defined as the service of any [item][7] at some level of organization for a higher level of organization (Zweers 1979: 423).

or:

> The terms role, biological role, biological function, biological meaning and function are taken to be equivalent. We prefer the use of the term "role". A role is defined as the service of a member for the system(s) to which the member belongs (Zweers 1979: 423).

Hence, a system is characterized by three types of parameters: structure, action and role.

According to Zweers the difference between actions and roles is relative to the level of investigation:

> If we take a particular structure or action parameter of a member (*i.e.* subsystem), this parameter can become a role parameter if the member is taken as the system being studied, without the member's relation with the higher system. Thus the shortening of the lingual muscle will be considered as an action of the lingual system, but the same shortening is considered to be a role as soon as the lingual muscle itself is taken as the system selected for investigation apart from and instead of the lingual system (Zweers 1979: 424)

---

[7] Zweers uses the term 'element'.

In this quote, Zweers depicts the shortening of the lingual muscle as an action of the lingual system, that becomes a role of the lingual muscle if the latter is taken as the item under study. In other words: according to Zweers the shortening of the muscle is seen as a role if it serves as the activity to be explained in an explanation of how a certain item (the lingual muscle) works, and the shortening of the muscle is seen as an action if it is used to explain how a system (the lingual system) that contains that item (the lingual muscle) performs its role. In my account, whether or not a certain change is called an activity (function$_1$) or role (function$_2$) does not depend on the way it is used in explanations. The shortening of a muscle is an activity (function$_1$), because in saying that a muscle shortens one does not say much about its effect on a larger system. This activity can serve as the activity to be explained in an explanation of how the lingual muscle works. It can also be used to explain the activity of a system of which it is a part. For instance, it might be used to explain how the lingual system works. As part of the latter explanation one will have to attribute a role to the muscle and its activity, for instance to move a certain bone, in a certain direction.

3.5 Tinbergen (1963)

The classical treatment of function as survival value in biology is Tinbergen's "On aims and methods of Ethology" (1963). The paper is a classic in ethology, but it has almost been ignored in the philosophy of biology. As far as I know, Horan (1989) is the only philosophical paper that discusses Tinbergen's article. Tinbergen distinguishes four problems of biology: causation, survival value (or function), ontogeny and evolution (p. 411). He observes (p. 417) that many biologists in his time are suspicious of explanations in terms of survival value. They hold that exact, experimental methods do not apply to the study of survival and that the assessment of survival value must, therefore, detoriate into unscientific guesswork. Tinbergen argues that this suspicion is undeserved. He appears to have three arguments. First, he lists many examples of studies demonstrating survival value as good as anyone could wish. His other two arguments are of a more philosophical nature: they concern the nature of the relation under investigation and the existence of a reliable method to demonstrate that relation. Tinbergen emphasizes that both physiology and the study of survival value are concerned with cause-effect relationships.

> The only difference is that the physiologists looks back in time, whereas the student of survival value, so-to-speak, looks "forward in time"; he follows events after the observable process has occurred (Tinbergen 1963: 418).

In Tinbergen's view, this difference between the study of physiological causes and of survival value is, just "an accident of human perception". In the case of physiological studies, due to our constitution, the effect is the thing that is easily observed and the cause is something to be discovered. However, in ethology the thing that is observed most easily (a certain behaviour) is the cause and the effect (its survival value) is the thing to be discovered. Tinbergen’s third argument concerns the method of investigation (Actually, Tinbergen himself appears to be unaware of the difference between the two arguments, he seems to equate cause-effect relations and relations that can be observed repeatedly). He is convinced that the undeserved suspicion of the study of survival value is due “to a confusion of the study of natural selection with that of survival” (p. 418). Whereas evolutionary studies of selection deal with unique past events that cannot be observed repeatedly and, hence, can never be subjected to experimental proof, the study of survival value (like physiological and developmental studies) is concerned with a flow of events that can be observed repeatedly and which thus can be subjected to experiments. The survival value of a certain morphological or behavioural element, therefore, can be established in a reliable way:

> The method to demonstrate survival value of any attribute of an animal is to try whether or not the animal would be worse off if deprived of this attribute” (Tinbergen 1963: 419).

This is easy with morphological items but if the attribute is a behavioural pattern it is often rather difficult to create an experimental organism that lacks just the attribute to be studied. This difficulty may be overcome by systematic comparisons of the success of animals at times when they do show a certain behaviour with the success at times when they do not, and by the use of dummies, such as plastic sticklebacks.

Chapter 4: The classical attempts

4.1 Introduction

In chapter 2 I described the practice of reasoning about functions in functional biology. I aim to explain what these kinds of reasoning add to our knowledge. I focus on the kind of reasoning which I have called ‘design explanation’. In this chapter I examine the now classical attempts of Carl Hempel (1959)[1] and Ernest Nagel (1961, 1977) to analyse the meaning of ‘function’ and to account for the explanatory force of reasoning that appeals to function. These attempts constitute the point of departure of many later discussions.

Both Hempel and Nagel employ an inferential theory of explanation. On this theory explanations work by showing that the phenomenon to be explained is to be expected in virtue of the explanatory facts. Applied to reasoning about functions, this means that a function attribution should allow one to infer the presence of the item to which the function is attributed if such an attribution is to be explanatory. According to Hempel, so-called ‘functional analyses’ aim to show that an organism is in such conditions that the trait under study has an effect that satisfies a need. Hempel argues that because, in general, the trait under study is not the only trait that may satisfy the need, functional analyses do not suffice to derive the conclusion that the trait under study is to be expected. Functional analyses are therefore explanatory only in the limited sense that they allow one to infer that one of the elements must be present of an ill-defined class of traits that may satisfy the need. Nagel gives another analysis. He argues that given the form of organization of a certain organism the presence of a certain item is a necessary condition for a certain function to be performed. Hence, given the fact that a certain function is performed, we may derive the conclusion that the corresponding item is present. Such derivations constitute, therefore, valid functional explanations.

I shall argue that both attempts are unsatisfactory. Hempel appears to be concerned with design explanations that explain the need to perform a certain causal role. He is right that the possibility of functional equivalents precludes the conclusion that a particular kind of item must be present. However, he draws the wrong conclusion from this observation. His conclusion is that appeals to need are explanatory to a limited extend only. I argue that the proper conclusion is that the inferential theory fails to account for what is learned from a design explanation.

Nagel argues that the problem of functional equivalents does not occur if the relevant conditions and the function in question are sufficiently detailed. I argue that this move is unsatisfac-

---

[1] I will quote this paper from its reprint in Hempel (1965), p. 297-330.

tory for several reasons. First, because none of the premises of the resulting argument is lawlike this move does not safeguard the explanatory character of appeals to function on the inferential account. Second, this moves deprives functional explanations from providing an important insight: namely that different structure is different animals might be seen as different solutions to the same problem. Last (but not least!), it misrepresents the structure of explanations as these are put forward by functional biologist.

4.2 Hempel (1959)

This section consists of roughly two parts: one part (4.2.1, 4.2.2) deals with Hempel's analysis of 'function' and 'functional analysis'. In section 4.2.1 I explain Hempel's analysis of the meaning of these notions. According to Hempel functional analyses aim to show that a certain activity or behavioural pattern satisfies a need. Functions are defined as traits satisfying needs. In section 4.2.2 I argue that Hempel is wrong in this identification of having a function and satisfying a need. The second part (section 4.2.3, 4.2.4 and 4.2.5) deals with Hempel's appraisal of the explanatory force of functional analyses. In section 4.2.3 I give some examples of studies that Hempel would label 'functional analysis'. In section 4.2.4 I explain Hempel's attempt to account for the insights such studies provide. On Hempel's account they have a very weak explanatory force but an important heuristic value. In section 4.2.5 I argue that analyses that show that a trait satisfies a need do have an explanatory character and that Hempel's fails to account for this character.

4.2.1 Hempel's account of 'function' and 'functional analysis'.

Hempel starts his discussion of functional explanation with the observation that it is often claimed that, in contrast to the physical sciences, biological, social and historical sciences cannot confine themselves to establishing causal or correlational connections. Proper understanding of the phenomena studied by these disciplines is supposed to require other types or methods (Hempel uses these words interchangeably) of explanations. One of the explanatory methods that has been developed for this purpose is the method of 'functional analysis'. Functional analysis is typically invoked to explain some recurrent activity or behavioural pattern in an individual or a group by appeal to its contribution to the preservation or development of the individual or the group in which this activity occurs. Hempel aims

> to examine the logical structure of functional analysis and its explanatory and predictive significance by means of a confrontation with the principal characteristics of the explanatory procedures used in the physical sciences (Hempel 1965: 297).

Hempel's main interest appears to be the use of functional analysis in the social sciences and his account of functional analysis owes much to a paper of the sociologist Robert Merton (1957). However, Hempel begins his discussion of functional analysis by considering a variant of the philosopher's standard example of a function attribution in biology:

> The heartbeat in vertebrates has the function of circulating blood through the organism (Hempel 1965: 305).

As he sees it, the meaning of this statement can not be expressed by replacing 'function' with 'effect', for this would make the production of heart sounds one of the functions of the heart, which it is obviously not. Hence, a first requirement for a philosophical theory of function is that it distinguishes between effects that are functions (such as circulating the blood) and effects that are side-effects (such as heart sounds). Hempel seeks this distinction in the fact that circulation, but not heart sounds, contributes to the satisfaction of certain requirements, the satisfaction of which is indispensable for the organism to remain in proper working order. Hempel proposes the following analysis of the foregoing function attribution:

> The heartbeat has the effect of circulating the blood, and this ensures the satisfaction of certain conditions (supply of nutriment and removal of waste) which are necessary for the proper working of the organism (Hempel 1965: 305).

More generally, functions are effects that satisfy needs. This suggests the following "basic pattern of functional analysis":

> The object of functional analysis is some "item" *i*, which is a relatively persistent trait or disposition (e.g., the beating of the heart) occurring in a system *s* (e.g., the body of a living vertebrate), and the analysis aims to show that *s* is in a state, or internal condition, c_i and in an environment representing certain external conditions, c_e such that under conditions c_i and c_e (jointly to be referred to as *c*) the trait *i* has effects which satisfy some "need" or "functional requirement" of *s*, i.e., a condition *n* which is necessary for the system's remaining in adequate, or effective, or proper, working order (Hempel 1965: 306)

In other words, a functional analysis is an attempt to show that in the conditions in which the organism lives the item in study has an effect that satisfies a need. Hempel says nothing about the relation between a functional analysis and a function attribution, but I take it that he takes it that a function attribution expresses the result of a functional analysis.

4.2.2 Why 'having a function' is not the same as 'satisfying a need'.

Hempel defines functions in terms of needs. He does not distinguish different kinds of functions. Moreover, he does not discuss any detailed example of a functional analysis in biological research. Nor does he give bibliographic references to such an example. This makes it difficult to determine what kind of study he has in mind when he talks of functional analysis and what kind(s) of function he wants to define in terms of needs. At first sight it seems that the kind of study he has in mind is a search for causal roles (the search for an answer to a type (2) question). Consider, for example, once more, Hempel's example of a function attribution:

> The heartbeat in vertebrates has the function of circulating blood through the organism (Hempel 1965: 305).

This example differs from function attributions in morphology in that the function of circulating the blood is attributed to an activity (the heartbeat), rather than to an item (the heart).[2] Apart from that it is clearly concerned with the causal role (function$_2$) of the heart in circulating the blood. Functional analysis might thus be seen as a search for causal roles. On Hempel's account, however, functional analysis does not merely aim to find out how a certain activity contributes to a complex activity or capacity. In addition, the functional analysis must show that the performance of the activity to which the activity under study contributes is (in its turn) a necessary condition for the organism to function adequately. To attribute the function of circulation to the heartbeat one must not only show that the heart contributes to circulating the blood by beating but also that the organism needs the circulation of the blood. This way of looking at attributions of causal roles does not conform to biological practice. One aim of functional biology is to explain how a certain organism is able to meet the requirements imposed on that organism by the way it is built / works / behaves and the environment in which it lives. Hence, functional biologists will often look for causal roles that help to explain an activity that needs to be done or a capacity that is needed by the organism. Yet, it is the fact that it helps to explain a certain activity or capacity that makes the causal role a causal role, and not the fact that that activity or capacity is needed. I have three arguments to support this claim that one should distinguish between satisfying a need and having a causal role (function$_2$).[3]

First, biologists are ready to talk of functions$_2$ (causal roles) even in cases in which the performance of this function is not needed to remain in proper working order. For example, the glandular hairs on the leafs of sundew are said to have the function to catch flies, even in circumstances in which sundews can survive without capturing prey.[4]

---

[2] Perhaps this is Hempel's way to mould the complex function attribution 'the heart contributes to circulating blood by beating' into the philosopher's standard form ('the function of ... is ...').

[3] George Williams's (1966) argument that one should distinguish between the needs a trait satisfies and its function as selected effect (function$_4$) was discussed in section 2.2.4.

[4] The fact that biologists are ready to say that the hairs of the sundew have the function to catch flies even in cases in which that function is not needed shows that Hempel's reading of this function attribution would be wrong but not that my reading is the right one. On my account the attribution of the function to catch flies to the hairs at the sundew's leaf is an attribution of a causal role. This interpretation explains the linguistic

Second, when searching for functions biologists often do not pay attention to the question whether or not the activities of an item satisfy a need. There is, for example, no mention of needs or demands in Harvey's (1628) account of the function of the heart (example 2.1 of chapter 2), neither in Miller's (1961) account of the function of the thymus (example 2.2).

Third, many design explanations explain the character of the item to which the causal role is attributed by appeal to that causal role without appeal to the needs satisfied by that causal role. For example, Schwenk's (1994) explanation of the form of the snake's tongue (example 2.3) appeals to the fact that the tongue has a causal role in trail-following, but not to the need to perform that causal role. He appeals to the fact that having a trail following role imposes demands on the tongue. The issue whether or not this function itself needs to be performed is irrelevant to this explanation. Hence, an account of the explanatory force of such explanations should not define functions in terms of needs.

These three arguments show that in order to attribute a function as a causal role (function$_2$) to an item it is sufficient to discover how that item contributes to an activity or capacity of a containing system and that one should distinguish between having a causal role (function$_2$) and satisfying a need.

4.2.3 Examples of functional analyses: the need to circulate oxygen

Although 'having a causal role' and 'satisfying a need' should not be identified (as I have shown in the preceding section) it is certainly the case that many studies in functional biology aim to show that the performance of a certain causal role (function$_2$) satisfies a need. Such studies aim for a design explanation that explains why it is useful to perform a certain task (that is a design explanation that answers a type (4a) question). Perhaps, it is this kind of study rather than the search for causal roles that Hempel had in mind when he talks of "functional analysis". In this section I give some examples of this kind of analysis, in the next two sections I will use these examples to show that Hempel fails to account for their explanatory force.

Outstanding examples of design explanations that explain the need to perform a certain task are explanations that are concerned with the need to circulate oxygen. The basics for such explanations were established by Krogh (1941). Krogh's work provoked a break through in respiratory biology. Krogh established that all oxygen transport ultimately relies on two kinds

---

behaviour of the biologists: functions as causal roles are determined by what an item does or is capable of doing rather than by the needs it satisfies. Proponents of an etiological reading of this function attribution would explain the biologists' behaviour by pointing out that the function is determined by what items of this kind *did* in the past (that accounts for the current presence of items of this kinds) rather than by their current needs. My arguments for reading this type of function attributions as attributions of causal roles, rather than as attributions of selected effects are given in chapter 7.

of physical process: diffusion and convection. The principles of diffusion are given by Fick's law of diffusion. This law states that the rate of diffusion of a gas is proportional to the gradient of partial pressure:

$$J = -D A \, \mathrm{d}P/\mathrm{d}x$$

In which:

J	the rate of diffusion (mm^3/s)
D	the diffusion coefficient ($\mathrm{mm}^2/\mathrm{atm}*\mathrm{s}$)
A	the surface area available for diffusion (mm^2)
P	the partial pressure of the diffusing gas (atm)
x	the distance of diffusion (mm)
$\mathrm{d}P/\mathrm{d}x$	the gradient of partial pressure (atm/mm)

For an organism to be able to survive and reproduce the oxygen supply must meet the demand. The oxygen supply at a certain point in the body of an organism is determined by the rate of diffusion. For an organism that has to rely on diffusion alone the relevant distance is that between the organs and the periphery. It follows from Fick's law that the rate of diffusion decreases with the distance if the concentration gradient remains the same. Hence, an organism that has to rely on diffusion alone will run into trouble if the distance between its organs and the periphery is too long. Krogh estimated that the radius of a hypothetical spherical organism living in water saturated with air cannot exceed 0.5 mm if it is to fill its need for oxygen by mere diffusion. Such an organism needs a system of convection in addition to diffusion. The system of blood circulation in Vertebrates provides such a system of convection. Other organisms employ other kinds of convection systems. Insects, for instance, transport oxygen by means of trachea (small tubes that circulate air) and sponges and coelenterates transport oxygen by means of water currents. All these systems satisfy the need for a system of convection in addition to diffusion.

Whereas in the above explanation the size of a "larger" organism explains the need for a circulatory system in such organisms, the absence of a circulatory system in its turn explains the small size of organisms that lack such a system. For example, McNeill Alexander (1979) argues that "flatworms are less than a millimetre thick because oxygen could not diffuse into them fast enough if they were thicker" (p. ii, see also p. 183). This conclusion is based again on a derivation using Fick's law of diffusion.

Another example concerns the respiratory pigments like haemoglobin and haemocyanin which are present in the blood of many animals. These pigments serve as oxygen-carriers: they bind the oxygen in the capillaries of the respiratory sites and release it in the capillaries of the organs. This function attribution answers a type (2) question (what is the causal role of the respiratory pigments?): it describes the causal role of the respiratory pigments in the circulatory system (this attribution helps to explain how the organism is able to circulate oxygen). It is appropriate not only to ask how respiratory pigments are able to perform this task (how is oxygen bonded, how is it released and how is this regulated?—type (3) questions), but also why the performance of this task is needed (why are oxygen-carriers needed?–type (4a) questions). The short answer to the latter question is that the solubility of oxygen in a simple saline solution is too low to carry enough oxygen to supply the tissues with oxygen at the required rate. McNeill Alexander (1979: 275-280) explains in more detail why the gastropod *Helix* needs a respiratory pigment. In order to do so, he calculates the rate at which the heart of *Helix* should pump the blood if the blood would not contain respiratory pigments. This calculation supports the conclusion that “the tissues could not be supplied with oxygen at the required rate unless the heart were larger or beat faster” (p. 276). The blood of *Helix*, however, is not a simple saline solution, but contains haemocyanin. Animals that carry oxygen by means of haemocyanin are able to carry $2^1/_2$–3 times as much oxygen as will dissolve in a physical solution. This suffices to meet the demand.

A fourth example is McNeill Alexander’s (1979: 357-259) design explanation of why intertidal polychaetes (for instance *Arenicola*) need gills, whereas earthworms can do without. Once again this explanation employs Fick’s law of diffusion. Earthworms and polychaetes both have a circulatory system. The distance between the superficial blood vessels and the air is about the same in earthworms and in polychaetes. McNeill Alexander calculates that “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” (p. 356). The thickest earthworms have diameters around 25 mm. Earthworms generally take their oxygen from the air. Polychaetes, however, take their oxygen from water. Oxygen diffuses much less fast through water than through air. According to Fick’s law and keeping all other things equal this would result in a rate of diffusion too low to meet the demands. Polychaetes solve this problem by irrigating their burrows. This keeps the partial pressure of oxygen high enough to maintain the required rate of diffusion. However, irrigation is impossible for intertidal species at low tide. As a result, keeping other things equal, the rate of diffusion would decrease. The gills solve this problem by increasing the surface area available for diffusion.

As I explained in section 4.2.1, Hempel describes a functional analysis as an attempt to show that in the conditions that apply to the organism in study the item in question has an effect that satisfies a need. This description applies to the examples above. So, let us see whether or not Hempel is able to account for the explanatory force of these examples.

4.2.4 Hempel's account of the scientific value of functional analyses

After having discussed the meaning of function attributions and the basic pattern of functional analysis, Hempel turns to an appraisal of the scientific value of such analyses. He observes that "functional analysis is widely considered as an *explanation* of the 'items' whose functions it studies" (p. 308). In his view proponents of functional analysis purport to explain the presence of a certain item by showing that it has some effects that satisfy a need. Hempel argues that the explanatory force of functional analyses is much more limited. This is due to the possibility of so-called 'functional equivalents', that is of different ways to satisfy a need or requirement. Hempel thinks of man made devices such as artificial hearts that might circulate the blood. Other examples of functional equivalents can be found in the examples above. I have mentioned three different ways to satisfy the need for a system of oxygen convection in addition to diffusion: blood circulation, trachea, and water currents. Further, both haemocyanin and haemoglobin may solve the need to carry oxygen.

According to Hempel, the possible existence of functional equivalents precludes the conclusion that a certain trait is present from the observation that a certain requirement is met. Consider the following pattern of explanation of an item (trait *i*) by functional analysis:

> (a) At *t*, *s* functions adequately in a setting of kind *c* (characterized by specific internal and external conditions)
>
> (b) *s* functions adequately in a setting of kind *c* only if a certain necessary condition, *n*, is satisfied
>
> (c) If trait *i* were present in *s* then, as an effect, condition *n* would be satisfied
>
> (d) (Hence), at *t*, trait *i* is present in *s* (Hempel 1965: 310)

In this pattern a description of the phenomenon to be explained (d) is derived from a combination of statements describing general laws (b and c) and a statement describing initial conditions (a), just as in a deductive-nomological explanation. However, in contrast with a deductive-nomological explanation, the conclusion (d) does not follow deductively from the premises (a-c), because it might well be that some trait *i*' different from *i* would suffice to satisfy need *n*. Conclusion (d) could be validly inferred only if (c) is replaced by (c"): 'requirement *n* can be met *only* if trait *i* were present in *s*'. In other words, in order to derive the conclusion that trait *i* is to be expected, trait *i* must not merely satisfy a need, it must be *indispensable* to satisfy that need. Hempel argues that his modified premise (c") is usually false. For example, an artificial pump can, perhaps, be used to pump the blood around. A functional analysis allows one only to derive the "very weak" (p. 313) conclusion that one of the several possible sufficient conditions is present. Therefore, the explanatory import of functional analysis is "limited to the precarious role" (p. 314) schematized in this pattern:

> (a) At *t*, *s* functions adequately in a setting of kind *c* (characterized by specific internal and external conditions)
>
> (b) *s* functions adequately in a setting of kind *c* only if a certain necessary condition, *n*, is satisfied
>
> (c') *I* is the class of empirically sufficient conditions for *n* in the context determined by *s* and *c*; and *I* is not empty
>
> (d') Some one of the items included in *I* is present in *s* at *t* (Hempel 1965: 313)

With respect to the predictive value of functional analysis Hempel observes that the possibility of functional equivalents limits the predictive power of functional analysis just as that possibility limits the explanatory power of functional analysis. Moreover, even the weak pattern given above can not readily be applied in prediction, for we do not know whether or not premise (a) (the organisms functions adequately) applies at some future time. To use this schema in prediction one should add a hypothesis to the effect that within certain limits the system under analysis will develop the means to satisfy its future needs (Hempel calls this "a hypothesis of self-regulation"). Hempel emphasizes that this hypothesis must be stated in an objectively testable form. In sum:

> [The] explanatory force [of functional analysis] is rather limited; in particular it does not provide an explanation of why a particular item *i* rather than some functional equivalent of it occurs in system *s*. And the predictive significance of functional analysis is practically nil—except in those cases where suitable hypothesis of self-regulation can be established (Hempel 1965: 324).

This does not mean that such analyses do not add to our knowledge. In Hempel's view their scientific value is to be sought in their contribution to the process of discovery rather than in their contribution to explanation or prediction:

> Functional studies in biology have been aimed at showing, for example, how in different species, specific homeostatic and regenerative processes contribute to the maintenance and development of the living organism; and they have gone on (i) to examine more and more precisely the nature and limits of those processes (this amounts basically to establishing various specific empirical hypotheses or laws of self-regulation), and (ii) to explore the underlying physiological or physiochemical mechanisms, and the laws governing them, in an effort to achieve a more thorough theoretical understanding of the phenomena at hand (Hempel 1965: 329/30)

Hence, on Hempel's account, analyses which show that a certain organism needs to perform a certain task (that is design explanations of the utility to perform a certain task) have a very weak explanatory value. Their main upshot is that they prompt biologists to study mechanisms of self-regulation. This analysis fails to do justice to the insights provided by such analyses in biology, as I shall show now.

4.2.5 Why Hempel’s account fails

In the previous section I described Hempel’s attempt to account for the explanatory force of analyses that show that a certain organism needs to perform a certain task (that is of a certain type of design explanation) by means of the inferential theory of explanation. On this theory analyses that show that a certain trait satisfies a need are explanatory if and only if such analyses allow us to infer the presence of the item that satisfies the need from the observation that the need is met. Hempel rightly observes that on this theory the explanatory force of such an analysis is rather weak because of the existence of functional equivalents. In this sections I shall use the examples of section 4.2.2. to show that Hempel draws the wrong conclusion from this observation. He draws the conclusion that functional analyses really have a very weak explanatory power. The proper conclusion is that the inferential theory fails to make sense of the explanatory power of functional analyses (that is of design explanation).

Consider Krogh’s analysis of the need for a circulatory system. It does not report newly discovered phenomena or laws. Nor does it yield any insights into mechanisms. What does this analysis add to our knowledge? The main insights provided by this study are insights in (1) how the need for a circulatory system is connected to the size of an organism, its activity and its environment, (2) how blood circulation, trachea and water currents are all solutions to the same problem and (3) how the need for a circulatory system relates to Fick’s law of diffusion. McNeill Alexander’s studies provide insights in (1) how flatness is connected to the absence of a circulatory system, the activity of the flat organism and the state of its environment, (2) how flatness is related to Fick’s law of diffusion, (3) how the presence of respiratory pigments is connected to the physical properties of the blood, the nature of the heart, the activity of the organism, and the environment, (4) how one difference between earthworms and polychaetes is related to their different environments.

Because of these insights biologists think of these analyses as explanatory. Hempel, however, finds himself enforced to deny the explanatory character of these analyses on the ground that they do not allow us to infer the presence of a particular item and dismisses the feeling that they are explanatory as an illusion of hindsight:

> The information typically provided by a functional analysis of an item *i* affords neither deductively nor inductively adequate grounds for expecting *i* rather than one of its alternatives. The impression that a functional analysis does provide such grounds, and thus explains the occurrence of *i*, is no doubt at least partly due to the benefit of the hindsight: when we seek to explain an item *i* we presumably know already that *i* has occurred (Hempel 1965: 313).

This puts the cart before the horse. Biologists are well aware of the existence of functional equivalents and they know that design explanations do not provide grounds for expecting one functional equivalent rather than another. Hence, it is not the illusion that a design explanation provides grounds for expecting a certain item that makes them think of design explanations as explanatory. However, the intuition that functional analyses are explanatory in combination with the awareness that functional analyses do not provide grounds for expecting a certain item should make philosophers think that providing grounds for expecting a certain item is not an adequate account of what makes an account explanatory. Let me emphasize that this is not a linguistic point. My point is not that Hempel's account fails as a conceptual analysis of what biologists call explanation, but rather that his account fails to account for the fact that design explanations add to our knowledge. Let us now see whether Nagel's account fairs better.

4.3 Nagel

4.3.1 Nagel's account of the meaning of function attributions

Nagel's focus is the question whether or not the use of teleological language in biology and the rejection of teleological explanation in the physical sciences entails the autonomy of biology from the physical sciences. Teleological statements are characterized by the occurrence of

> such typical locutions as 'the function of', 'the purpose of', 'for the sake of' and the like—more generally, the occurrence of expressions signifying a means-end nexus (Nagel 1961: 403).

An example of such a teleological statement is the following function attribution:

> the function of chlorophyll in plants is to enable plants to perform photosynthesis (i.e., to form starch from carbon dioxide and water in the presence of sunlight)[5] (Nagel 1961: 403).

A second example:

> The function of leucocytes in human blood is to defend the body against foreign microorganisms (Nagel 1961: 405).

Nagel argues that teleological statements can be translated without any loss of asserted content into non-teleological ones. However, he is not very clear about the form this translation is supposed to take. In fact, he suggests at least four different schemes.

At p. 403 Nagel (1961) states that teleological statements are "telescoped arguments" which when unpacked explain the presence of a certain item (chlorophyll, leukocytes) by showing that the presence of this item is a necessary condition for the occurrence of an activity (photosynthe-

---

[5] To put the record straight: biologists distinguish between photosynthesis, which is the production of organic carbon (sugar) from inorganic molecules in the presence of light, and the synthesis of starch from the sugars produced by photosynthesis. The first process takes place in the presence of light in the green parts of the plant. The second process does not depend on light and occurs also in storage organs such as the potato tuber.

sis, defence against micro-organisms) performed by the organisms that have the item (plants, humans). Such unpacked explanations are valid explanations in accordance with the deductive-nomological model. The attribution of the function of photosynthesis to chlorophyll, for instance, could be unpacked as follows:

> When supplied with water, carbon dioxide, and sunlight, plants produce starch;
>
> If plants have no chlorophyll, even though they have water, carbon dioxide, and sunlight, they do not manufacture starch;
>
> hence, plants contain chlorophyll (Nagel 1961: 403).

More generally, function attributions of the form

> The function of A in system S with organization C is to enable S in environment E to engage in process P

can be reformulated as:

> Every system S with organization C and in environment E engages in process P;
>
> if S with organization C and in environment E does not have A, then S does not engage in P;
>
> hence, S with organization C must have A (Nagel 1961: 403).

Nagel does not always stick to the idea that function attributions are telescoped arguments. Sometimes, he merely holds that teleological statements of the form “the function of *i* in organisms of type *s* is to enable *f*” can be translated into non-teleological statements of the form “*i* is a necessary condition of *f*” without loss of “asserted content”. For example at p. 405 of *The Structure of Science* Nagel states that

> the initial, unexpanded statement about chlorophyll appears to assert nothing that is not asserted by ‘Plants perform photosynthesis only if they contain chlorophyll,’ or alternatively by ‘A necessary condition for the occurrence of photosynthesis is the presence of chlorophyll’ (Nagel 1961: 405).

On this view the functions of an item are processes or activities in which that item takes part and for which the existence of that item is a necessary condition. A teleological statement expresses consequences (effects) and is equivalent to a causal (“non-teleological”) statement asserting conditions necessary for the performance of a certain activity. Function attributions, on this view, are explanatory not because they are telescoped functional explanations but because, they may be used (in their non-teleological reformulation) in functional explanations as the lawlike premise.

At page 406 Nagel raises the problem why functions are attributed to the parts of living organisms but not to the parts of physical systems such as gases and solar systems. He is interested in this problem because he suspects that some biologists will use this difference in the application of function attributions to argue that there is a difference between function attributions and non-teleological statements that is not captured by Nagel’s foregoing analysis. Those biologists might bring up the following counter-example against Nagel’s account: “the function of the pressure varying inversely with the volume is to keep the temperature constant”. On Nagel’s account this statement would be a valid function attribution, since, according to Boyle’s law, the temperature of a changing volume of gas remains constant only if the pressure varies inversely with the volume. However, few physicists, if any, would be prepared to accept this statement. Those biologists might subsequently claim that it is the “goal-directed” character of organisms which makes teleological statements appropriate to biology and that the fact that teleological explanations are usually advanced only in connection with “goal-directed” systems affects the claim that teleological explanations are equivalent to non-teleological ones.

Nagel’s answer to this problem is basically this: indeed, function attributions presuppose that the system under consideration is goal-directed and that the function ascribed to an item contributes to the realization of some goal for which the system is directively organized.[6] However, according to Nagel, the notions of ‘goal’ and ‘goal-directed system’ can be analyzed into non-teleological terms. The presupposition, therefore, does not affect the thesis that every teleological explanation is translatable into a non-teleological one.

Nagel’s example of a goal directed system is the system that maintains the temperature of the human body within a narrow range around 37°C. This goal is achieved by means of a number of mechanisms. The most important source of heat is the energy lost in oxidative metabolism. The basal metabolic rate is influenced, among others, by a hormone produced in the thyroid gland. Under many circumstances the wasted heat is sufficient to keep us warm enough. If one gets too cold shivering may provide an additional source of heat. The main way to loose heat is via the skin. The heat radiated through the skin depends on the quantity of blood flowing through the peripheral blood-vessels, a quantity which is regulated by dilation or contraction of this vessels. Sweating is another important mechanism to get rid of excess heat. The adrenal gland produces a hormone that affects internal metabolism and shivering.

According to Nagel, goal-directed systems are characterized by the fact that they are “directively organized”. Roughly speaking, a system *S* (for instance the system that maintains the inner temperature) is directively organized (“goal-directed”) with respect to a certain goal-state *G* (e.g. the temperature falls within a specified range) during a certain time if during that time:

---

[6] This is most clearly stated in his "Teleology Revisited" (1977):

> A functional statement of the form: a function of item *i* in system *S* and environment *E* is *F*, presupposes (though it may not imply) that *S* is goal-directed to *some* goal *G*, to the realization or maintenance of which *F* contributes (Nagel 1977: 297)

> (1) *S* consist of a set of related parts or processes (the state variables) the states of which jointly determine the occurrence of *G* in S (some relevant state variables in the temperature example are: the states of the peripheral blood vessels, the thyroid gland, and the adrenal glands);
>
> (2) a change in one of the state variables (a primary variation) will take the system out of the goal-state if unaccompanied by changes in the other state variables;
>
> (3) the parts are so related that as a matter of fact a primary variation results in changes in the other state variables in such a manner that the goal-state is maintained (adaptive variation).

If this analysis of goal directed systems as directively organized systems succeeds, Nagel has shown that the observation that function attributions are usually advanced only in connection with goal-directed systems does not support the thesis of the impossibility of the translation of teleological language into non-teleological language. However, Nagel's discussion of goal-directed system does not answer the original objection that the translation of function attributions into statements of the form "*i* is a necessary condition of *f*" is incomplete. In fact, Nagel's discussion supports this objection and suggests a revision along the following lines:

> A statement of the form "a function of item *i* in systems of type *s* is to enable *f*" can be reformulated as:
>
> (1) systems of type s are directively organized with respect to at least one goal
>
> (2) in systems of type *s*, item *i* is necessary to perform f
>
> (3) the occurrence of *f* contributes (at least under appropriate circumstances) to the realization or maintenance of at least one of the goals with respect to which systems of type s are directively organized.

On this view function attributions are not equivalent to teleological explanations, because on this revised account function attributions assert more than the corresponding teleological explanations do (namely that the function contributes to a goal state). However, neither in chapter 12 of *The Structure of Science*, where Nagel discusses "the structure of teleological explanations" in biology, nor in "Teleology Revisited" is this possible revision and its relation to functional explanation explicitly discussed. Nagel largely restricts himself to the observations that when biologists attribute functions they are usually interested in how a certain item contributes to the maintenance of the organism (p. 408, 422) and that they usually presuppose that the system they are interested in is goal-directed (p. 421). He fails to draw the consequences from this observation for his analysis of function attributions and functional explanation. The item's contribution to the maintenance of a goal-state plays no role whatsoever in Nagel's account of functional explanation.

However, in chapter 14 of *The Structure of Science*, where Nagel discusses "functionalism in social science" he talks in yet a different mode. For example, at page 525 he distinguishes as one meaning of the term 'function' in social science "the sense that has occupied us in chapter 12". In this sense

> the function of some item signifies the contribution it makes (or is capable of making under appropriate circumstances) toward the *maintenance* of some stated characteristic or condition in a given system to which that item is assumed to belong (Nagel 1961: 525)

This quote suggests not merely that function attributions presuppose that the function contributes to the maintenance of some goal-state or other (as in the third translation scheme), but rather that the goal-state is explicitly mentioned in function attributions. Moreover, as Lehman (1965a: 9) points out, in this part of his book Nagel seems to have forgotten his earlier contention that items are necessary for their function . Indeed, Nagel warns his readers that they should not think, that functional items are indispensable:

> One of the functions of the thyroid glands in the human body is to help preserve the internal temperature of the organism. However, this is also one of the functions of the adrenal glands, so that in this respect there are at least two organs in the body that perform (or are capable of performing) a similar function. Accordingly although the maintenance of a steady internal temperature may be indispensable for the survival of human organisms, it would be an obvious blunder to conclude that since the thyroid glands contribute to this maintenance they are for this reason indispensable for the continuance of human life. Indeed, there are human beings who, as a consequence of surgical intervention, do not have thyroid glands, but nevertheless remain alive (Nagel 1961: 533).

Subsequently, he scorns Malinowski for committing this fallacy:

> This point has not been consistently recognized by functionalists. For example, Malinowski argued that because the function of myth is to strengthen tradition by attributing to it a supernatural origin, “myth is, therefore, an indispensable ingredient of all culture” (Malinowski 1948: 146). However, although one may grant, if only for the sake of the argument, the role Malinowski ascribes to myth in strengthening traditions, as well as his tacit claim as to the indispensability of tradition in all societies for the persistence of their cultures, his conclusion is nevertheless a *non sequitur*. For he transfers without warrant the admitted indispensability of *tradition*, to a *particular means or instrument* that happens to be employed in certain societies for sustaining tradition (Nagel 1961: 533).

Note that Nagel’s example of a function attribution in biology (“the function of the thyroid is to help preserve the internal temperature”) only states that a certain item contributes to a certain goal. Malinowski’s function attribution (“the function of myth is to strengthen tradition by attributing to it a supernatural origin”) on the other hand states both the goal (maintaining the tradition) to which the item (myth) contributes and the manner in which that item contributes to the maintenance of the goal state (by giving it a supernatural origin). These remarks suggests that function attributions might by more or less complete. They are complete if they specify both the goal-state and the means by which the item contributes to the realization or maintenance of that state. They are incomplete if they specify only one of these two. Nagel’s main example in chapter 12 (the function of chlorophyll in plants is to enable plants to perform photosynthesis) now emerges as an incomplete function attribution that specifies the means but not the goals. Nagel's discussion of function attributions in chapter 14 suggests the following translation scheme for complete function attributions:

> A statement of the form "A/the function of item *i* in systems of type *s* is to contribute to the maintenance of goal *g* by means of *m*" (e.g. "the function of myth is to strengthen tradition by attributing to it a supernatural origin") can be reformulated as
>
> (1) systems of type *s* are directively organized with respect to *g*
>
> (2) in systems of type *s*, item *i* effects *m* (at least under appropriate circumstances)
>
> (3) in systems of type *s*, the occurrence of *m* contributes (at least under appropriate circumstances) to the realization of *g*.

Incomplete function attributions mention only the goal

> A statement of the form "A/the function of item *i* in system *s* is to contribute to the maintenance of goal *g*" (e.g. "the function of the thyroid in human bodies is to help humans to preserve the internal temperature") can be reformulated as
>
> (1) systems of type *s* are directively organized with respect to *g*
>
> (2) in systems of type *s*, item *i* contributes to the realization or maintenance of *g*

or the means

> A statement of the form "A/the function of item *i* in system *s* is to perform or enable *m*" (e.g. "the function of chlorophyll in plants is to enable photosynthesis") can be reformulated as
>
> (1) systems of type *s* are directively organized with respect to at least one goal state
>
> (2) in systems of type *s*, item *i* effects *m* (at least under appropriate circumstances)
>
> (3) in systems of type *s*, the occurrence of *m* contributes (at least under appropriate circumstances) to the realization of at least one of the goals with respect to which systems of type *s* are directively organized.

Nagel's account of functional explanations leaves the explanatory force of this kind of function attributions entirely unexplained.

In this section I have argued that Nagel provides at least four different analyses of function attributions, namely (1) function attributions as telescoped arguments, (2) function attributions as stating that the presence of a certain item is necessary condition to perform a certain activity, (3) function attributions as stating that a certain item is necessary for a certain effect which in its turn helps to maintain a certain goal-state, (4) function attributions as stating how a certain item contributes to the maintenance of a certain goal state. On the first three schemes function attributions imply that the item to which the function is attributed is necessary for its function. On Nagel's account of functional explanation, it is this implication (and only this implication) that is important in a functional explanation. In the next section I discuss the question whether or not such an analysis sufficiently accounts for what is achieved by appeals to function in explanations in functional biology.

4.3.2 Nagel's account of functional explanation

As I have emphasized in section 4.2.3 attributions of needs as they are used in functional biology typically allow for functional equivalents, that is they allow for different ways to meet a certain requirement. For example, many organisms satisfy the need for a system of convection by means of a system of blood circulation, insects circulate air and sponges and coelenterates circulate water. Haemoglobin and haemocyanin can both satisfy the need for a oxygen carrier and so on. I have also emphasized that such attributions have an important explanatory role. Nagel (1961, 1977) tries to account for the explanatory force of attributions of needs in terms of the inferential theory of explanation, just as Hempel (1959) does. On this account, explanations work by showing that the phenomenon to be explained is to be expected in virtue of the explanatory facts. As I have discussed in section 4.2.4 Hempel argues that the existence of functional equivalents prevents the derivation of the conclusion that a certain item is present from a functional analysis. In order to show that the phenomenon to be explained was to be expected the item under study must be shown to indispensable. Hempel rightly observes that, usually, this is not the case and draws the conclusion that analyses that appeal to needs are explanatory only in a very limited sense. Nagel on the other hand argues that, given a certain form of organization, the items of which the presence is explained by means of a functional explanation are, indeed, indispensable for their function and, more general, for that organism to survive. Analyses that appeal to functions are, therefore, really explanatory.

As I have discussed in section 4.3.1 according to Nagel (1961) functional explanations conform to the deductive-nomological model. They have the following form:

> Every system S with organization C and in environment E engages in process P;
>
> if S with organization C and in environment E does not have A, then S does not engage in P;
>
> hence, S with organization C must have A (Nagel 1961: 403).

For example:

> When supplied with water, carbon dioxide, and sunlight, plants produce starch;
>
> If plants have no chlorophyll, even though they have water, carbon dioxide, and sunlight, they do not manufacture starch;
>
> hence, plants contain chlorophyll (Nagel 1961: 403).

Nagel's reconstruction of functional explanations in "Teleology Revisited" (1977) is slightly different from that in *The Structure of Science* (1961). The conclusion of the explanation is now rendered as a statement about an individual system (e.g. a certain plant) during a certain period, rather than about a certain type of systems (e.g. plants) in general. The argument consists of three premises, rather than two:

> The explanatory premises for the assertion having the form "the item *i* occurs in *S* during a given period *t* and circumstances *E*"—for example, "during a stated period and given circumstances, chlorophyll is present in the specified green plant"—are as follows:
>
> (1) "During a stated period, the system *S* is in environment *E*" (for example, "during a stated period, a green plant is provided with water, carbon dioxide, and sunlight");
>
> (2) "During that period and in the stated circumstances, the system *S* does *F*" (e.g., "during the stated period, and when provided with water, carbon dioxide, and sunlight, the green plant performs photosynthesis")
>
> (3) "If during a given period *t* the system *S* is in environment *E*, then if *S* performs *F* the item *i* is present in *S*" (e.g., "if during a given period a green plant is provided with water, carbon dioxide, and sunlight, then if the plant performs photosynthesis the plant contains chlorophyll") (Nagel 1977: 300)

In this scheme:

> the first two premises are instantial statements, and the third is lawlike (Nagel 1977: 300)

The differences between these two schemes of functional explanation are not important for the criticisms I will bring up in section 4.3.3.

In regard to functional equivalents, Nagel observes in *The Structure of Science* that it has been objected against his analysis that the second premise of his explanatory scheme would be untenable if chlorophyll is not necessary to produce starch or if plants can maintain themselves without photosynthesis. He argues that this objection rests on a misunderstanding. Nagel concedes that it is logically possible that there are plants which do not produce starch or which produce starch in absence of chlorophyll. He even points out that there are in fact plants that survive without chlorophyll. However, the functional explanation of the occurrence of chlorophyll is not concerned with logical and physical possibilities but with a definite form of organization, the so-called "green plants". Given this form of organization there is no reason to question the need for chlorophyll:

> It is certainly *logically* possible that plants might maintain themselves without manufacturing starch, or that processes in living organisms might produce starch without requiring chlorophyll. Indeed, there are plants (the funguses) that can flourish without chlorophyll; and in general, there is more than one way of skinning a cat. On the other hand, the above teleological explanation of the occurrence of chlorophyll in plants is presumably concerned with living organisms having certain determinate forms of organization and definite modes of behavior—in short, with the so-called "green plants." Accordingly, although living organisms (plants as well as animals) capable of maintaining themselves without processes involving the

> operation of chlorophyll are both abstractely and physically possible, there appears to be no evidence whatever that in view of the limited capacities green plants possess as a consequence of their actual mode of organization, these organisms can live without chlorophyll (Nagel 1961: 404, emphasis in original).

To avoid such misunderstandings function attributions should detail the form of organization to which they apply:

> a teleological explanation must articulate with exactitude both the character of the end-product and the defining traits of the systems manifesting them, relative to which the indicated processes are supposedly indispensable (Nagel 1961: 404).

A similar argument is given in “Teleology Revisited” where Nagel discusses Hempel’s argument that the possibility of functional equivalents (such as artificial hearts that pump the blood around) deprives functional analyses of much of their purported explanatory power. Nagel argues that since physiologists are concerned with normal human beings, observations about humans with artificial hearts are not relevant:

> A convincing case can be made for the claim that in normal human beings—that is, in human bodies having the organs for which they are at presently genetically programmed—the heart *is* necessary for circulating blood; for in normal human beings there are in fact no alternative mechanisms for effecting the blood’s circulation. For physiologists seeking to explain how the blood is circulated in normal human bodies have discovered that human bodies have no organs other than the heart for performing that function. The observation that it *may* be (or actually *is*) physically possible to circulate blood by means of other mechanisms is doubtfully relevant to those investigations of how the blood is circulated in the human beings, upon which physiologists are embarked (Nagel 1977: 292).

He repeats his idea that the impression that there are functional equivalents is due to the imprecise way in which the system to which the function is attributed is specified:

> The denial of the claim that the heart is necessary for circulating the blood appears to derive part of its plausibility from the imprecise way in which the expression ‘human body’—and more generally the expression ‘the system S’ is usually specified (Nagel 1977: 292).

He adds to this that the problem of functional equivalents is well-known in the context of causal explanation. At first sight death may result from a plurality of causes (e.g. drowning, poison, gun shots and so on), none of which is in itself indispensable for its effect. However, if the condition of the corpse is specified in a more precise way only one type of cause remains for each type of death and the occurrence of this cause is a necessary condition for all effects of this type. Similarly, in the case of functional explanation the problem of functional equivalents should be solved by a more precise description of the system in which the item functions. In the example of the function of the heart the system in which the heart has a function should be described as a “normal human being” rather than as a “human being”. Given this form of organization the presence of a heart is a necessary condition to pump the blood around.

In sum, Nagel’s way of dealing with the problem of functional equivalents in the context of explanations that appeal to functions appears to be: (1) to emphasize that the scope of the premise stating that the presence of a certain item is indispensable to perform a certain function is restricted to those organisms (“systems”) that indeed have that item, and (2) to deny on that ground the relevance of observations about possible or actual organisms that lack that item. In the next section I argue that this way of dealing with the problem of functional equivalents not only confuses attributions of needs with attributions of functions, but also prevents the relevant premise of its lawlikeness. As a result, Nagel’s analysis fails to account for the explanatory force of attributions of needs.

4.3.3 Physical possibilities

As I have shown in the previous section, Nagel repeatedly suggests that functional biologists are concerned with what is the case in actual existing organisms rather than with logical or physical possibilities. This is a severe misunderstanding. It is certainly true that functional biologists are not concerned with *logical* possibilities, but *physical* possibilities play an important role in functional biology, just as in any other natural science. More specifically, the point of many design explanations is to show that the existing form is physically possible whereas certain alternative forms are not. Consider, for example, Krogh’s study of the need for a circulatory system. As I have discussed in section 4.2.3, Krogh showed with the help of Fick’s law of diffusion that organisms larger than a certain size that rely on diffusion alone are physically impossible. Another example is McNeill Alexander’s explanation of why *Helix* needs respiratory pigments (example 3 in section 4.2.3). He uses the laws of physical chemistry to establish (1) that organisms that are built in the way in which *Helix* is built but which do not have respiratory pigments are physically impossible because they cannot satisfy their need for oxygen, and (2) that the presence of haemocyanin solves this problem.

This concern with physical possibilities is a crucial difference between attributions of needs and attributions of causal roles (function$_2$ attributions). Attributions of causal roles are generalizations about how a certain item is actually used in a certain group of organisms. Attributions of needs on the other hand are lawlike. They do not merely tell us what is the case (in a certain group of organisms), but what is physically possible and what not (given the conditions stated in the attribution). For example (example 2.3 of chapter 2), Schwenk’s attribution of a trail-following role to the tongues of all snakes and certain lineages of lizards (an attribution of a causal role) is a generalization about how the tongue is used in those organisms. It does not allow any conclusion about yet unknown species. Perhaps there are yet undiscovered species of snakes that have stopped using their tongue in this way; perhaps there are lineages of snakes that have never developed this habit. If such a species were discovered, the statement that the tongue has a trail-following role would still apply to the species to which it applies now. Schwenk's observation that in order to use chemosensory tropotaxis those organisms must be able to sense simultaneously the chemical stimuli at two points (an attribution of a need), however, is not concerned with what is actually the case in those organisms, but with the conditions under which chemosensory tropotaxis is physically possible (or not). From this attribution one may derive valid conclusions about unknown species. For instance, one may derive the conclusion that unknown species that use their tongue in chemosensory tropotaxis must have either a forked tongue or a functional equivalent device (that is another device that enables them to sense simultaneously the chemical stimuli at two sides of the body). If a species were found that does not meet this requirement but still uses its tongue in chemosensory tropotaxis this would give reason to doubt Schwenk's attribution of a need. Hence, although observations about possible or actual organisms that lack the item to which the function$_2$ is attributed are irrelevant to attributions of causal roles such observations are, *pace* Nagel, highly relevant to attributions of needs.

Because attributions of causal roles are nothing but generalizations about how a certain item is used or about how a certain causal role is performed, an attribution of a causal role in itself does not explain the item to which the causal role is attributed. To explain an item by appeal to its causal role one needs a lawlike statement in addition to the attribution of a causal role. Often this lawlike statement takes the form of an attribution of a need. For example, to explain the forkedness of the snake's tongue, Schwenk uses the attribution of the need to sense simultaneously the chemical stimuli at two points in addition to the attribution of a trail-following role. This explanation can be expressed by means of the following train of thought:

> (1) The tongues of snakes have a role in chemosensory tropotaxis
>
> (2) Chemosensory tropotaxis is physically 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$_2$[7] the tongues of snakes are forked

Note that what is explained by appeal to the trail following role is not the presence of the item to which the function is attributed (the tongue) but its forked character. Both Hempel and Nagel erroneously assume that function attributions are used to explain the presence of the item to which the function is attributed.

---

[7] Recall that 'why$_2$' is short for 'why it is useful that' (see section 2.3.2).

More generally, the train of thought of many design explanations that appeal to the demands imposed on an item by the causal roles (function$_2$) that item has (that is design explanation that address a type 4b question), can be expressed by means of the following scheme of explanation:

(1) In *s*-organisms: item *i* performs causal role *f*

(2) The performance of causal role *f* is physically possible only if requirement *n* is satisfied

(3) In *s*-organisms: if item *i* has character *c* then requirement *n* is satisfied

----

(4) That's why$_2$ in *s*-organisms item *i* has character *c.*

The train of thought represented by this scheme is not an argument in the sense in which the term 'argument' is used in logic: the explanatory statements are not intended to support a conclusion, but to explain the phenomenon described in the last statement. The first statement in this scheme is an attribution of a causal role (function$_2$), the second one an attribution of a need and the third one states how the need is satisfied in a certain group of organisms. The first statement is a generalization about the causal role of a certain item in a certain group of organisms. The third statement is a generalization about the way in which a need imposed by that causal role is satisfied in that group of organisms. The group of organisms mentioned in the first and the third statement is not necessarily a systematic or phylogenetic category. An enumeration of the individuals to which these statements apply would work too. The second statement is lawlike. Instead of an attribution of a need biologists often use a weaker kind of lawlike statement to explain the character of an item, namely the statement 'in order to perform function *f* it is useful if condition *n* is satisfied'.

Design explanations that appeal to the need to perform a certain causal role (design explanations that address a type 4a question), too, consists of a lawlike statement in addition to the function attribution. Many of them follow the following scheme:

(1) s-organisms are in condition c_n

(2) in condition c_n: organisms are able to survive and reproduce only if causal role *f* is performed

(3) in s-organisms: item *i* performs causal role *f*

----

(4) That's why$_2$ s-organisms have item *i*

An example is Krogh's (1941) explanation of why$_2$ vertebrates have a circulatory system:

(1) The distance between the inner organs of Vertebrates and their periphery is more than 1 mm

(2) If the distance between the inner organs and the periphery is more than 1 mm an organism is viable only if it has a system of convection in addition to diffusion

(3) The circulatory system of Vertebrates provides a system of convection

---

(4) That's why$_2$ vertebrates have a circulatory system.

In this explanation again the attribution of a causal role (function$_2$) (3) is combined with a lawlike statement (2) in order to generate the explanation. The latter statement, an attribution of a need, expresses what is physical possible.

A philosophical theory of explanation in functional biology should (among other things) account for the explanatory character of trains of thought of this kind. On the inferential account this train of thought would be explanatory if and only if the explanatory statements give us reason to expect the phenomenon to be explained. In order to do so the scheme must represent a valid argument, the premises of which must be true and at least one of the premises must be lawlike. Because of the possibility of functional equivalents (ad. (3)) the schemes of the previous paragraph are invalid. For that reason the inferential account fails to make sense of the explanatory character of this kind of design explanations.

As I discussed in the previous section, Nagel attempts to save the inferential account from the above objection by observing that given a certain form of organization there is in fact only one kind of item that performs the function. In his view a function attribution in itself explains the presence of the item to which the function is attributed. Given an appropriate specification of both the function and the system to which the function attribution applies there is only one type of item in the world that performs this function. For that reason one may infer the presence of the item to which the function is attributed from the observation that the function is performed. This move saves the validity of the argument at the cost of the lawlikeness of the general statement (which is the second premise in the scheme of *The Structure of Science* and the third in "Teleology Revisited"). This is most clearly illustrated by means of his own example: the explanation of the presence of chlorophyll.

Nagel does not discuss this example in any detail. Nor does he give bibliographic references. This makes it difficult to determine exactly what explanation he has in mind. Let us take a look at the history of the study of chlorophyll to see if we can make sense of his example. In section 2.2.2 I discussed several examples of attributions of causal roles. In all those cases one was looking for an unknown causal role of a known item. The case of chlorophyll is quite different. The causal role was known before the item. The history of the study of photosynthesis starts in 1772. In that year the famous English chemist Joseph Priestley reported to the

Royal Society that air rendered impure by a breathing mouse or a burning candle could be purified by a living plant. A few years later Jan Ingenhousz, a Dutch physician, discovered that this process of purification took place only in sunlight and only in the green parts of plants. In 1782 the Swiss Jean Senebier showed that the process depends on the presence of a gas he called ‘fixed air’ (carbon dioxide). Another Swiss researcher, Nicolas Théodore de Saussure, discovered the role of water in 1804 and established the following overall equation of the photosynthetic process:

carbondioxide + water + light --> organic matter + oxygen

This is essentially the overall equation as it is known now. In 1817 two French chemists, Pelletier and Caventou, managed to extract the green substance from the leaves. They called this substance ‘chlorophyll’. From this point in the history of biology it was known which substance performed photosynthesis. The structure of this substance became known only in the mid of the twentieth century.

Biologists do not think of the attribution of the role to perform photosynthesis to chlorophyll as an explanation of the presence of chlorophyll. When biologists ask the question ‘why do plants have chlorophyll?’ (e.g. Mauzerall 1977, Seely 1977) they have in mind very specific type (4b) questions about the structure of chlorophyll. This structure was established from degradation studies by H. Fischer in Germany in 1940. R.B. Woodward confirmed this analysis in 1960 by a complete synthesis of the molecule. The chlorophyll molecule contains a porphyrin ‘head’ and a phyttol ‘tail’. The porphyrin head is made of a tetrapyrole ring containing a magnesium atom. There are several forms of chlorophyll and there are many other pigments (such as carotenoids and phycobillins) involved in photosynthesis. However, only two forms of chlorophyll are involved in the photochemical process itself: chlorophyll a is used in all forms that employ oxygen and bacteriochlorophyll is used in photosynthetic bacteria. All other pigments serve as light harvesting pigments: they absorb light in a variety of wavelengths and transfer the excitation energy to chlorophyll a or bacteriochlorophyll. One way of explaining why plants have chlorophyll is answering the question ‘why do plants use chlorophyll a rather than one of the available alternatives (carotenoids and phycobillins) as the first step in the photochemical process?’. Part of the answer is that chlorophyll absorbs in the low energy end of the visible spectrum. Radiation with less energy does not produce chemical reactions. Radiation with more energy is harvested by the other pigments. For physical reasons, a harvesting pigment needs to dump its energy at an energy level lower than the energy of the light it absorbs. For that reason it is useful if the receiver pigment absorbs energy at a level that is as low as possible. Other issues in this area of research concern questions about the presence of specific organic groups and the question why magnesium rather than some other metal is trapped in the tetrapyrole ring.

Hence, when biologists study the type (4b) question ‘why do plants have chlorophyll’ they try to explain why$_2$ the photochemical reaction is performed by a molecule which has the structure chlorophyll has by appealing to the usefulness of having a certain structure in performing the causal role in photosynthesis. Nagel’s scheme does not represent this practice. When he talks about an explanation of chlorophyll he has something much less interesting in mind, namely a derivation of the conclusion that chlorophyll is present from the observation that photosynthesis is performed together with the general statement that all green plants that perform photosynthesis have chlorophyll (this is the second premise in the scheme of *The Structure of Science* and the third in “Teleology Revisited”). This view of functional explanation not only fails to do justice to the practice of functional biology, it also fails to reveal the explanatory force of the function attribution on the inferential account. To count as an explanation on the inferential account the general statement must be a lawlike. On Nagel’s account, however, this statement is only a generalization about what is the case in a certain group of plants.

The trouble is the definition of ‘green plants’. Nagel is not quite clear how this form of organization is defined. ‘Green plants’ might simply mean those plants that are green. Because the green colour of green plants is due to the presence of chlorophyll, the statement “all green plants perform photosynthesis only if they have chlorophyll” is a truism on this reading, rather than a law. Moreover, many photosynthetic algae are not green but red, brown or purple. On this reading the attribution of the causal role to perform photosynthesis would not apply to this group of organisms. Yet, according to biologists the causal role of chlorophyll in these algae is to enable photosynthesis.

On the other hand, Nagel could try to identify ‘green plants’ with a certain taxonomic group, such as the kingdom of plants. Folk biology tends to divide living organisms into two kinds: plants (including algae, fungi and perhaps bacteria) and animals. Biologists are more sophisticated and distinguish *at least* six kingdoms: (1) eubacteria, (2) archaebacteria, (3) protista, (4) plants, (5) fungi, (6) animals. There is a lively discussion on this subject (the protista group is highly problematic) but it is undisputed that the fungi are not to be included in the kingdom of plants and, for that reason, it would seem to help Nagel to relativize the attribution to this group in order to avoid functional equivalents. However, some of the organisms included in the kingdom of plants (no matter its exact definition) do not need chlorophyll: parasitic forms such as the Birdsnest orchid and all species of the Broomworth family do not perform photosynthesis but get their energy from sugars taken from their hosts. Saprophytic forms such as *Monotropa* also do not have chlorophyll but extract sugars from leaf litter, with the help of funguses. Furthermore there are certain groups of eubacteria (which on no account belong to the kingdom of plants) that use chlorophyll to perform photosynthesis. The same problem occurs if one would try to define green plants as a phylogenetic group

4.4 Conclusions

Hempel and Nagel are concerned to account for both function attributions and design explanations. I have argued against Hempel’s meaning analysis that ‘having a function’ is not the same as ‘satisfying a need’ and against Nagel’s meaning analysis that ‘having a function’ is not the same as ‘being needed’. Although causal roles are not *defined* in terms of needs, Hempel and Nagel were right in thinking that the kind of reasoning which functional biologists call ‘functional explanation’ (and which I have called ‘design explanaton’) appeals to needs and requirements. However, the attempts of Hempel and Nagel to account for such explanations on the inferential theory of explanation fail.

My diagnosis of Hempel’s and Nagel’s failure is as follows. Both Hempel and Nagel are of opinion that the existence of functional equivalents would prevent an account from being explanatory. This poses the following dilemma: either there exist real functional equivalents in which case one should deny that appeals to needs are explanatory or appeals to function are really explanatory in which case one should deny the existence of real functional equivalents. Hempel takes the first horn of this dilemma, Nagel the second one. Hempel’s move is not acceptable because it fails to account for what is achieved by a design explanation. Nagel’s move does not work because it saves the validity of an appeal to needs at the cost of the law-likeness of that appeal. But even if Nagel’s attempt to downplay the existence of functional equivalents had been succesful his account would have had severe problems. One of the things one may learn from a design explanation is precisely that two or more different structures or behaviours are functional equivalent. Krogh’s explanation, for example, yields the insight that blood circulation, trachea and water currents might all provide different solutions to different problems. An account of explanation that denies the existence of functional equivalents must deny these insights and is for that reason not acceptable. That is, any account of explanation in functional biology should allow for functional equivalents. Apparently there is something wrong with the inferential conception of explanation.

Chapter 5: The causal role theory

5.0 Introduction to chapter 5 to 8

In the previous chapter I argued that the attempts of Hempel and Nagel to account for the explanatory force of function attributions in terms of the inferential theory of explanation fail.

In chapter 5 to 7 I discuss more recent approaches to function: the causal role theory, the survival value approach, and the etiological account. In chapter 8 I present my own account. In chapter 5 to 7 I aim to evaluate the extend to which the different approaches account for the explanatory force of appeals to the different kinds of function as they are used in biology, especially in the disciplines known as functional morphology and ethology (see chapter 2). It will turn out that each of these approaches is concerned with only one type of function: the causal role theory is concerned with function as causal role (function$_2$), the survival value approach with function as survival value (function$_3$), and the etiological account with function as selected effect (function$_4$). As I mentioned in chapter 1, understanding the practice of biological explanation is not always the aim with which the theories I discuss were originally proposed. The proponents of the different theories often fail to realize that there are different notions of function and present, for instance, examples of attributions of causal roles (function$_2$) to support an analysis that applies only to function as survival value (function$_3$). It is important to distinguish the extend to which a philosophical theory of function captures the *meaning* of the different notions of function (as these are used in biology) from the extend to which that theory accounts for the *explanatory use* of appeals to the different kinds of function.

Chapter 5 is concerned with Cummins's (1975, 1983) causal role theory. I argue that this theory gives a correct analysis of the meaning of attributions of causal roles (function$_2$) and also that this theory accounts for one important explanatory use of attributions of causal roles, namely their use in what I call ‘capacity explanations’. There are, however, other important explanatory uses of attributions of causal roles that are left unexplained, namely their use in design explanations and their use in evolutionary selection explanations.

Chapter 6 is concerned with the survival value approach. In the first part of this chapter I show that the survival value approach is on the right track as an analysis of the *meaning* of claims about survival value. I elaborate my own account of the meaning of such claims (presented in section 2.2.3) and show that the objections that have been brought up against other proposals within the survival value approach do not apply to my account. In the second part of chapter 6 I argue that the current proposals with the survival value approach fail to give a satisfactory account of the *explanatory use* of appeals to survival value.

Chapter 7 deals with the etiological account. I argue that appeals to function in explanations in functional biology concern function as causal role (function$_2$) and function as survival value (function$_3$). As the etiological theory does not apply to these kinds of function, this theory is irrelevant to my subject.

The main upshot of my discussion in chapter 5 to 7 is this:

- Cummins provides a satisfactory account of the notion of function as causal role (function$_2$);
- I do provide a satisfactory account of the notion of function as survival value (function$_3$);
- Cummins provides a satisfactory account of appeals to causal roles in capacity explanations;
- an account of appeals to causal roles and to survival value in design explanation is lacking;
- an account of appeals to causal roles in evolutionary explanations is lacking.

In chapter 8 I attempt to work out a theory that should fill the lacuna.

5.1 Introduction to chapter 5

Hempel and Nagel tried to account for the explanatory force of function attributions in terms of the inferential theory of explanation. On this theory explanations show that the phenomenon to be explained was to be expected in virtue of the explaining facts. Most of the more recent attempts employ the causal theory of explanation. On this theory explanations reveal the mechanisms and processes that bring about the phenomenon to be explained.

One of the first attempts to account for the explanatory force of function attributions in terms of the causal theory of explanation is that of Cummins (1975) (see also Cummins 1983, chapter 1 and 2). Cummins rejects the idea that functional explanations explain the presence of a certain item by attributing a function to that item. According to him, explanations that explain the presence of a certain item must appeal to factors that cause that item to be present. As the performance of a function is an effect of an item’s presence rather than a cause, one cannot explain the presence of that item by appeal to its function. On Cummins’s account, functional explanations explain the capacities of a system of which the item in study is a part. For example, the appeal to the heart’s function to pump blood serves to explain the capacity of the system of heart and blood-vessels to circulate oxygen, nutrients and waste.

In the section 5.2 I describe Cummins’s criticism of previous account of functional explanations and his own theory of functionel explanation in more detail. In section 5.3 I show by means of an example that functional biologists indeed employ a kind of explanation (capacity explanation) that conforms to Cummins’s theory. In addition, I show by means of several examples that function attributions in Cummins’s sense are used in other kinds of explanation as well. Theses uses are ignored or denied by Cummins. In section 5.4 I show that Cummins’s criticisms do not apply to the use of function attributions in design explanations. In section 5.5 I argue that Cummins’s arguments against the use of attributions of causal roles in evolutionary explanations do not hold water. In section 5.6 I discuss some objections brought up against Cummins's view by other philosophers. Finally, in section 5.7 I draw some conclusions.

5.2 Cummins's account of function and functional explanation

In this section I relate Cummins's criticism of the classical view on attempts to account for functional explanations on the inferential theory of explanation (5.2.1), Cummins's own account of functional explanation (5.2.2), and Cummins's account of function attribution (5.2.3).

5.2.1 Cummins's criticism of previous accounts of functional explanation

Cummins starts his "Functional Analysis" (1975) with the observation that all accounts of functional explanation from Hempel (1959) and Nagel (1961) onward have proceeded under two undefended assumptions, namely (1) that the point of a function attribution is to explain the presence of the item to which the function is attributed, and (2) that functions are a special kind of effects on a containing system[1] (for example those effects that satisfy a need of an organism, or those effects that contribute to a certain goal state of an organism). Note, that the first assumption concerns the *explanatory use* of function attributions, and the second their *meaning*. Taken together these assumptions yield the view that function attributions explain why a certain item is present in a certain system by pointing out that that item is present because it has a certain kind of effect on that system. Given this view the main task of a philosophical theory of function is to characterize the kind of effect that are functions.

Cummins rejects both assumptions. In regard to the first assumption he maintains that explanations of the presence of an item should appeal to causally relevant factors. Function attributions are concerned with the effects of an item. Since the effects of an item are causally irrelevant to the presence of that item one may not appeal to functions to explain the presence of that item.

> The problem is [...] that to "explain" the presence of the heart in vertebrates by appeal to what the heart *does* is to "explain" its presence by appeal to factors that are causally irrelevant to its presence. Even if it were possible, as Nagel claimed, to *deduce* the presence of chlorophyll from the occurrence of photosynthesis, this would fail to *explain* the presence of chlorophyll in green plants in just the way deducing the presence and height of a building from the existence and length of its shadow would fail to explain why the building is there and has the height it does. This is not because all explanation is causal explanation: it is not. But to explain in the presence of a naturally occurring structure or physical process—to explain

---

[1] A containing system is a system of which the item to which the function is attributed is a part.

> why it is there, why such a thing exists in the place (system, context) it does—this does require specifying factors that causally determine the appearance of that structure or process (Cummins 1975: 745/6, italics in original).

In regard to the second assumption, Cummins argues that it is impossible to describe the kind of effects that are functions adequately. As a first attempt one may equate the functions of a part of a system with those effects that contribute to a function of that system. This system itself can be seen as a part of a more encompassing system and so on. The problem is that the chain of encompassing systems will have to end somewhere (for example with the organism) and the definition does not allow us to attribute functions to that final system. As a second attempt one may equate the function of a part of an organism with that part's contribution to that organism's health or survival. This definition has, at least, two problems. First, there are cases in which the exercise of a function is unfavourable to health and survival (for example, after mating many male spiders are eaten by their mate). Second, some healthy effects do not count as functions (for example, adrenaline secretion may effect loss of weight but this effect does not count as a function of adrenaline secretion). A third attempt may equate functions with effects that contribute to the organism's capacity to contribute to the survival of its species. According to Cummins, this attempt is too limited as we can easily imagine cases in which the exercise of a certain function would not contribute to the survival of the species. For example, if flying ceased to contribute to the pigeon's capacity to maintain its species we would still say that propulsion during flight is a function of the pigeon's wing.[2]

5.2.2 Cummins's account of functional explanation

Cummins presents his own account of functional explanation in section 3 of his "Functional Analysis" (1975: 758-761).and in the first two chapters of his The *Nature of Psychological Explanation* (1983: 1-27). Central to this account is the idea that functional explanation is a distinctive style of explanation that does not conform to the covering law model. This kind of explanation is used to explain complex capacities.

Cummins starts his account of functional explanation by contrasting two main strategies of explanation: subsumption and analysis.[3] Explanations by subsumption explain changes of state in a system as the effect of certain previous changes by appeal to causal laws. An example is the explanation of the increase in pressure in a certain container with gas as the effect of a change in

---

[2] Cummins does not discuss definitions that equate functions with those effects that contribute to an organisms inclusive fittness. The counterexample against the third defintion would apply to this definition too.

[3] Cummins's (1983: 1-27) account of the difference between subsumption and analysis differs from Cummins (1975: 758-761). In this paragraph I follow his 1983 account.

volume by appeal to Boyle's law. Analytical explanations, on the other hand, are not concerned with changes but with properties. They explain how the properties of a certain system are brought about in terms of the properties of the components of that system and their mode of organization. An example is the explanation of the temperature of a certain gas by the law that the temperature of a gas is the average mean kinetic energy of the molecules in the gas. Cummins calls a law that specifies how a property is embodied in a certain type of system an 'instantiation law'. An analysis of a system into parts is called a 'componential analysis'.[4] Analytical explanations often include another kind of analysis in addition to the componential analysis, namely an analysis of the *property* to be explained in terms of the properties of the parts. Cummins calls such an analysis a 'property analysis' if the property to be explained is non-dispositional, and 'functional analysis' if the property to be explained is a disposition or a capacity.[5] Functional analysis proceeds by analyzing a capacity into a number of other capacities which add up, flow chart style, to the capacity to be explained.[6]

An analysis might appeal to properties or components which themselves require analysis. A componential analysis must eventually terminate in what Cummins calls 'nomic attributions'. Nomic attributions are fundamental lawlike statements to the effect that all components of a certain kind manifest a certain property (an example is the law of inertia in classical mechanics).[7] A functional analysis must eventually terminate in dispositions which are explainable via componential analysis without further appeal to subcapacities. A functional analysis explains a capacity only if the explaining capacities at this elementary level are indeed instantiated by the system: functional analysis is the first stage of a functional explanation, instantiation is the second (Cummins 1983: 31).

According to Cummins, functional analyses are explanatory because they show us how a complex capacity is brought about by the co-ordinated activity of simple parts. The power, and, hence, the suitability of this style of explanation depends on the extend to which the explaining capacities are simpler than and different from the capacity to be explained, and on the relative complexity of the organization attributed to the system. This latter feature (complexity of organization) correlates with the first two: if the gap between the explaining capacities and the capacity to be explained is greater one needs a more complex organization to bridge that gap (Cummins 1975: 764, 1983: 30).

---

[4] Cummins also uses the terms 'system analysis' and 'compositional analysis'.

[5] Cummins (1983) uses the terms 'disposition' and 'capacity' as stylistic variants.

[6] Note, that Cummins's 'functional analysis' refers to a different activity as Hempel's. Hempel refers to the analysis of needs, Cummins to the analysis of a complex capacity into subcapacities.

[7] According to Cummins such nomics attributions require justification but not explanation.

5.2.3 Cummins's account of function attributions

According to Cummins, function attributions are best viewed as the products of functional analyses. The functions of an item are those capacities of an item which feature in a functional explanation of some complex capacity.

> In the context of science, to ascribe a function to something is to ascribe a capacity to it that is singled out by its role in an analysis of some capacity of a containing system. When a capacity of a containing system is appropriately explained via analysis, the analyzing capacities emerge as functions (Cummins 1983: 28) (see also Cummins 1975: 765).

Hence, it is the applicability of a certain strategy of explanation (viz. explanation by functional analysis) that justifies the use of the term 'function', rather than the kind of effect singled out. For example, the capacity of the heart to pump blood counts as a function because the capacity of an organism to circulate oxygen, waste and nutrients is properly explained by appeal to (among other things) the capacity of the heart to pump blood. (In section 5.3.1 I will elaborate on this example.) As there is no functional explanation that appeals to the heart's capacity to produce sounds, the production of heart sounds does not count as a function.

5.2.4 Summary of Cummins's account

Thus, on Cummins's view, (i) functional explanations explain a complex capacity of a certain system by appeal to more simple capacities of the parts of that system, (ii) such accounts are explanatory on a causal account of explanation because they show us how a complex capacity is brought about by a system of simple subsystems, (iii) the function of a part of a system is a capacity of that part which features in a functional explanation of a capacity of that system, and, hence, (iv) the question whether or not a certain function attribution is justified comes down to the question whether or not a functional explanation is suitable. In the next section I discuss the applicability of this account to explanations in functional biology.

5.3 Evaluation of Cummins's account

What insights in explanation in functional biology does Cummins's account provide? In section 5.3.1 I show by means of an example that Cummins's account of functional *explanation* applies to an important kind of explanation in functional biology, namely capacity explanation. In section 5.3.2 I show by means of examples that Cummins's account of function *attribution* accounts for one way in which attributions of causal role are used in explanations (namely their use in capacity explanations), but leaves several other uses out of sight (namely their use in design explanations and their use in selection explanations). It also neglects the explanatory use of attributions of survival value.

5.3.1 Functional explanation

According to Cummins, functional analysis is widely employed in functional biology. Biologists typically explain the biologically significant capacities of the organisms they study by analyzing that organism into a number of systems (such as the circulatory system, the digestive system, and the musculoskeletal system). Each of this systems is defined by its characteristic capacities. These capacities in turn are analyzed into capacities of the components and so on. Ideally, this strategy is carried on until a level is reached at which the explaining capacities are sufficiently simple to explain them in terms of the physical and chemical characteristics of the components (Cummins 1975:760/1, 1983: 29).

Cummins does not discuss any detailed example of functional explanation in biology, but it is not difficult to find explanations that conform to his analysis in textbooks on functional biology. In section 2.3.2 I have labelled such explanations ‘capacity explanations’. For example, explanations of an organism’s capacity to circulate oxygen typically start by observing that the circulatory system consists of, say, three parts: a heart, the blood, and a system of blood-vessels. Next, they point out the causal role (function$_2$) of each of this parts in maintaining the circulation: the blood carries oxygen, the heart pumps the blood around (the heart is the source of energy for blood movement) and the vessels are arranged in such a manner that blood is transported from the heart to the lungs (where it is aerated) via another part of the heart to the organs (where oxygen is released) back to the first part of the heart. The capacities of these parts are in turn explained in terms of the capacities of their subparts. For example, the capacity of the heart to pump the blood around is explained in terms of its internal structure, its ability to contract, its rhythmicity and the nervous control. The capacity of the blood to carry oxygen is explained by the presence of haemoglobin, a respiratory pigment that has the capacity to carry oxygen. The system of blood-vessels is further divided into arteries, veins and capillaries. And so on, until a level is reached at which the capacities of the parts are explained in terms of their physical and chemical properties, without appealing to capacities again.[8]

For example, the capacity of the heart muscle (and more generally the capacity of any muscle) to contract is explained by the sliding filament theory, first proposed in 1954 by H.E. Huxley and J. Hanson and independently by A.F. Huxley and R. Niedergerke. According to this theory the change in muscle length is caused by two kinds of filaments (thick filaments and thin filaments) that slide past each other during contraction. The mechanism of contraction can be explained in molecular terms. The thick filaments are composed mainly of the protein myosin; the thin filaments are composed mainly of another protein: actin. The thick filament consists of a bundle of myosin molecules. A myosin molecule is an elongated, club-

---

[8] Explanations of the capacity to circulate oxygen that match this rough sketch can be found in any textbook on functional biology, see, for example Johansen (1977) or Keeton and Gould (1993: 843-860).

shaped structure with a linear tail and a globular head. Due to this structure, myosin molecules aggregate into filaments in which the molecules are arranged in two opposite directions, with clusters of the globular heads at the distal ends and their linear tails overlapping. The globular heads form a series of cross bridges to the actin filament. The cross bridges act as levers that pull the actin filaments towards each other during contraction. Contraction at this level is a kind of ratchet mechanism. After being activated by binding one ATP molecule (the source of energy for contraction), a myosin head binds to a special receptor site at the actin molecule. The energy for the binding is provided by the splitting of ATP in ADP and phosphate. The release of these two derivates induces a change in the conformation of the myosin head due to which the actin filaments are pulled together. As the result of the bending the myosin disconnects from the actin and becomes free to start a new cycle. Due to the arrangement of the myosin molecules (in opposite directions) the actin filaments are pulled towards each other and the muscle shortens. Hence, the capacity of the muscles to contract is explained by (1) the arrangement of the thick and thin filaments, (2) the structure and arrangement of the molecules of which the filaments are composed, (3) changes in binding and in conformation of the molecules involved, (4) the release of ATP. At this level of explanation no further appeal to subcapacities is needed.[9]

This example indicates that Cummins has drawn attention to an important kind of explanation in functional biology, namely capacity explanation. This kind of explanation was completely ignored in the philosophy of science at the time Cummins's article appeared. Cummins account details the way in which capacity explanations are construed and brings to light the explanatory force of this kind of explanation (on a causal account): capacity explanations show us how complex capacities are brought about by the operation of systems of subsystems that are ultimately chemical or physical in nature.

5.3.2 Attributions of causal roles

In the previous section I discussed Cummins's account of functional *explanation*. I argued that this account reveals the explanatory force of an important kind of functional explanation, namely capacity explanation. In this section I evaluate Cummins's account of function *attributions*. In regard to this subject Cummins argues (i) that the assumption that the point of a function attribution is to explain the presence of the item to which the function is attributed is wrong (see section 5.2.1 above), and (ii) that functions are singled out by their role in capacity explanations (see section 5.2.3 above). As Cummins distinguishes between function and survival value (see Cummins 1975: 749-751) I will assume that these claims are meant to apply to

---

[9] This account of muscle contractility is derived from Huxley (1965) and Keeton & Gould (1993: 1076-1081); for a more elaborated discussion of the quest for an explanation of the heart's capacity to contract see Robinson (1986).

function as causal role (function$_2$) only. In this section I argue, by means of examples, that Cummins is right that attributions of causal roles are singled out by their role in capacity explanations, but that capacity explanation is often not the main point of an attribution of a causal role. Quite often causal roles are attributed as part of an attempt to explain the design or the evolution of the item to which the causal role is attributed. I will use the examples of attributions of causal roles I brought up in section 2.2.2: Harvey's claim that the principle causal role of the heart is to propagate blood (rather that to produce heat and life), Miller's study of the causal role of the thymus, and Schwenk's explanation of the snake's forked tongue. If the point of an attribution of a causal role is to explain a capacity of a containing system one would expect that research into causal roles is guided by a (type 3) question of the form 'how are such and such organism able to perform such and such activity?'. However, in these three examples the questions that guide research are questions of the type 'why do such and such organisms have such and such type of item?'. Cummins does not pay attention to the use of attributions of causal role in answering these kind of questions.

The causal role of the heart

As I explained in section 2.2.2, Harvey's (1628) claim that pumping blood is the function of the heart is part of an attempt to establish the thesis that the blood circulates continuously through the body. His main argument for this thesis concerns the quantity of blood passing through the heart. In addition he argues that the thesis of a continuous circulation would explain a wide variety of phenomena by showing "for what cause they are made" (p. 133). Phenomena that would be "very hard for any one to explain by any other way" (p. 133). One such phenomenon is the beating of the heart. Another is the difference in thickness between the muscular wall of the left and that of the right ventricle of the heart. The wall of the left ventricle is much thicker than the wall of the right one. On the hypothesis that the main function of the heart is to pump the blood around, this difference is easily explained. Because, the left ventricle has to pump the blood through the whole body it needs more power, and hence more muscles, than the right ventricle which has to pump the blood only through the lungs (p. 124).

In other words, Harvey argues that attributing the function to pump blood to the heart helps to answer the following questions 'why does the heart beat?' and 'why does the heart have the structure it has?'. No doubt, he singles out pumping because pumping explains the activity in which he is interested (circulation), but at this stage of his argument his point is that the assumption that the heart has the causal role to pump the blood around helps to explain the structure and activity of the heart. In this example, the function attribution serves as a first step in a design explanation of the structure and activity of the item to which the function is attributed.

In defence of Cummins's position one might try to dismiss my reference to Harvey as irrelevant.[10] "You aim to explain the current explanatory practice in biology", so the argument would go, "but the kind of explanation Harvey employed to explain the structure and action of the heart by appeal to its function is teleological in a way that is not acceptable today". To a certain extend I am sympathetic to this response. On the theory of explanation Harvey employed function attributions are explanatory because they identify the purpose for which the item to which the function is attributed was brought about. Today this is not acceptable as an account of why function attributions are explanatory. However, Harvey's function attribution is still accepted today, just as his explanation of the difference in thickness between the ventricles by appeal to this function. This can be checked in most textbooks on functional morphology. For example, Johansen (1977) observes that in birds and mammals

> the structural design of the left and the right ventricles reflects their function as a volume pump and a pressure pump respectively (Johansen 1977: 388).

Among the differences in structure which are explained by the different causal roles of the ventricles is the difference in the thickness of the walls (I have discussed this explanation in section 2.3.4.). This shows that Harvey's explanations are today accepted as explanatory, despite the fact that his concept of explanation is not. Our task is to show how Harvey's explanation makes sense on our canons even if we do not accept the kind of teleology he accepted. Cummins's account is of no help here.

The causal role of the thymus

It is clear that in the second example of section 2.2.2 (the causal role of the thymus) the point of the function attribution is to make sense of the thymus, not to explain some capacity. It is true that the function which is finally attributed to the thymus (the function to initiate the differentiation of T-lymphocytes) is singled out by its role in an explanation of the development of the immune system, but Miller's (1961) studies did not aim to explain the development of the immune system. Miller did not even know that the immune system develops. He aimed to find out how the thymus contributes to our physiology. The existence of the thymus was a complete riddle because it appeared to have no effect at all on the organism's physiology. Part of this riddle was solved by showing that and how the thymus contributed to the development of the immune system.

---

[10] In a similar vein, Neander (1991a) rejects the argument against the etiological approach that Harvey talked of functions without having heard of natural selection on the ground that "Harvey obviously did not have natural selection in mind when he proclaimed the function of the heart but that does not show that modern biologists do not have it in mind" (p. 176).

Hence, the question that guided research was the question ‘why do many vertebrates have a thymus?’. Miller’s experiments provide a partial answer to this question: the thymus has a causal role in the development of the immune system. This answer situates the thymus in the organization of the machinery by which vertebrates meet their needs. This answer is partial because it does not tell us why this task is performed by a specialized organ. It lasted until the 1980s before Langman (1989) came up with a design explanation that answers this question (I have outlined this explanation in section 2.2.3).

The snake’s forked tongue

In the third example of section 2.2.2 (the snake’s forked tongue) the function attribution (the tongue has a trail following role) is singled out by its role in an explanation of the capacity to find prey and mates. Yet, as the title of Schwenk’s (1994) paper (“Why Snakes have Forked Tongues”) indicates, the insight that the tongue has a trail-following role does not only serve to explain the snake’s capacity to find prey and mates, but also to explain why the tongue is forked. As I discussed the forking is explained at two levels. At the organismal level the forking is explained by means of a design explanation that appeals to the requirements imposed on the tongue by the mechanism of trail-following. At the historical level the increase of forking is explained by an evolutionary selection explanation that appeals to the fact that the causal role to follow trails was performed more efficiently by those individuals that had tongues that were forked more than average.

Conclusion

As I said in the introduction to this subsection, if the main point of an attribution of a causal role is to explain a complex capacity, one would expect that research into functions is guided by a question of the type ‘how are such and such organism able to perform such and such activity?’. However, in all three examples the question that guides research is not a question about a capacity but a question about an item. Harvey was interested in the structure and activity of the heart, Miller in the causal role of the thymus and Schwenk in the forked character of the tongue. As a first step in answering this question the researchers attribute a causal role to the item in which they are interested. This attribution situates the item in a system that performs a certain task. It tells us how that item contributes to that task. Subsequently, this hypothesis about the place of the item in the way in which the organism is organized is used to explain the presence or character of that item. This can be done at two levels. In a design explanation it is shown that the item performs its causal role better if it has the character it has (rather than if it had some other character). In an evolutionary explanation the character of the item is explained by showing how the causal role of that item influenced the evolution of that item. This means that in addition to the use of attributions of causal role in capacity explanations, there are at least two other uses of attributions of causal roles, namely (i) the use of attributions of causal role in design explanations, and (ii) the use of attributions of causal role in evolutionary explanations. Cummins appears to ignore this. As I said in section 5.2.1, Cummins argues that attributions of causal role cannot be used in explanations of the presence of the item to which the causal role is attributed. In the next sections I will discuss whether his arguments apply to the use of attributions of causal roles in respectively design explanations (section 5.4) and evolutionary explanations (section 5.5).

5.4 Design explanations and causality.

As I said before, Cummins (1975: 746) maintains that explanations of the presence of a naturally occurring item or physical process must appeal to causally relevant factors. Because, the performance of a causal role is the effect of an item’s presence (and not its cause) one may not appeal to the causal role of a certain item to explain the presence of that item.

In a certain sense (of ‘explain the presence of an item’), design explanations explain the presence of the item or behaviour to which the function is attributed. As I showed previously, design explanations are taken to explain why certain organisms have a certain kind of item or perform a certain kind of behaviour. Schwenk, for example, explains why snakes have a forked tongue (example 2.3 of section 2.2.2) and Krogh explains why larger organism have a circulatory system (see section 4.2.3).[11]

However, design explanations do not explain the presence of an item in the sense in which Cummins uses the phrase ‘explain the presence of an item’. Cummins seems to refer to explanations that explain how a certain kind of item became present in a certain place in the course of time. Design explanations do not explain the presence of an item or behaviour in that sense, they explain how a certain item or behaviour is useful (synchronically) not how it came about in the course of time (diachronically). This justifies the conclusion that Cummins’s thesis, that explanations of the presence of an item must appeal to causally relevant factors, cannot be used to argue against appeals to causal roles and survival values in design explanations. Such explanations do not pretend to explain the presence of an item in the way in which Cummins uses that phrase.

Cummins (1975: 746) explicitly states that there might be explanations that do not explain the presence of an item and that such explanations need not be of a causal nature. As I discussed in section 5.2.2, Cummins (1975: 758-763, 1983: chapter 1) distinguishes two kinds of explanations: explanations that explain changes (explanations by subsumption) and explanations that explain properties (analytical explanations). He tends to restrict the term ‘causal ex-

---

[11] See also section 2.3.2.

planation' to explanations of the first kind. However, both kinds of explanations are causal in the sense in which I use that word: explanations by subsumption show us how a certain change is brought about by preceding changes, analytical explanation show us how a certain property of an item or behaviour is brought about by the action of its components. Design explanations belong to neither of Cummins's categories: they do not explain changes and they do not analyze a thing into components. They are not of a causal nature because they are not concerned with the mechanisms that bring about a certain change, property, state or ability. Instead they show us why a certain item, behaviour or character is useful to its bearers. One of the main challenges of a theory of design explanation is to explain how this kind of reasoning can be explanatory.

5.5 Attributions of causal roles in selection explanations

In section 5.4 and 5.5 I deal with Cummins's arguments against the use of attributions of causal roles to explain the presence of the item to which the causal role is attributed. I discuss the question whether these arguments apply to the appeal to causal roles in design explanations respectively evolutionary explanations. As I said before, Cummins maintains that appeals to causal roles to explain the presence of a certain kind of item are to be rejected on the ground that causal roles are causally irrelevant to the presence of that item. In regard to design explanations I argued (in section 4.5) that design explanations are not meant to explain the presence of the item to which the function is attributed (in Cummins's sense of 'explain the presence'). In regard to appeals to causal roles in evolutionary explanations I shall argue (in this section) (i) that such explanations do explain the presence of certain kind of item, and (ii) that such explanations appeal to causally relevant factors. In section 5.5.1 I expose my own views on this subject. In section 5.5.2 I show that Cummins does not provide valid arguments against this view. In section 5.5.3 I draw the conclusion that there are no objections to appeal to causal roles to explain the evolution of the item to which the causal role is attributed.

5.5.1 A causal account of appeal to causal roles in evolutionary explanations

For a start, consider the historical part of Schwenk's explanation of why snakes have a forked tongue. At the historical level Schwenk explains the presence of forked tongues in currently living snakes as the result of a series of past changes at the level of the population. These changes resulted in a gradual increase of the forking of the tongue in the course of evolution. Schwenk explains this increase by appeal to selection to perform the trail-following role of that item more efficiently. This explanation assumes that in an ancestral population of snakes with moderately forked tongues certain mutations occurred as the result of which the tongues of the mutant organisms were more highly forked than average. Due to this increased forking the tongue of those mutants performed its trail-following role more efficiently than average. This increased the relative fitness of those mutant organisms. As a result, the mutation spread through the population and the average forking of the tongue in that lineage increased.

Schwenk, clearly, explains the presence of forked tongues in currently living snakes and he does so by appeal to the causal role of that item. The explanations says that one of the processes that gave rise to the phenomenon to be explained (the presence of highly forked tongues in currently living snakes), was selection for a better performance of the trail-following role. Thus, the fact that the tongue had a trail-following role is a positive causal factor for the increase of the forking in the course of the evolution: if the tongue had another role an increased forking would probably not have been advantageous and mutations with more than averagely forked tongues would probably not have spread through the population. Hence, this way of appealing to causal roles to explain the presence of a certain kind of item is acceptable on a causal view of explanation.

In order to avoid misunderstandings, let me emphasize two distinctions. The first distinction is a distinction between two levels at which processes may take place: the level of the organism (individual) and the level of the lineage (population). In Schwenk's explanation, it is the fact that the tongue had a trail-following role in *individuals in an ancestral population* that explains the appearance of highly forked tongues *in the lineage*.

The second distinction concerns two ways in which items may be grouped into kinds: on the basis of structural similarity and on the basis of homology. The concept of 'homology' is subject to discussion but for my purposes one should think of homologous items as items that have the same evolutionary origin.[12] For example, bird hearts are homologous to mammal hearts because if one traces their evolutionary history one would find that both these hearts originate as modifications from the heart of some reptile-like common ancestor. However, insect hearts are not homologous to vertebrate hearts because (to our best knowledge) the most recent common ancestor of insects and vertebrates did not have a heart (which shows that vertebrate hearts and insect hearts did not originate from one original). Although the tongues of all reptiles are homologous to each other, its forked character in the lineages that have a forked tongue is not homologous in all those lineages: Schwenk showed that this character evolved at least twice. In the context of evolutionary explanations two items in different organisms are seen as occurrences of the same item in the lineage if those items are homologous.

Hence, Schwenk explains the appearance of a certain kind of items (i.e. *items that have a certain structure*) in a lineage by appeal to the causal role of those items (i.e. *homologous items*) in past individuals of that lineage. Note, that in this kind of explanations the bridge between the

---

[12] A more precise definition can be given with help of Millikan's (1984: 23-25) notion of 'reproductively established family': two items are called 'homologous' if they belong to the same (higher-order) reproductively established family.

level of the individual and the level of the population, is provided by natural selection: an item changes (at the level of the lineage) due to heritable fitness differences between individuals as the result of heritable differences in the way in which that item performs its role in different individuals. The way in which an item performs its role is, therefore, a causally relevant factor in the evolution of the item to which the function is attributed. In the next section I discuss Cummins's views on this issue.

5.5.2 Cummins's view of causal roles and natural selection.

Cummins (1975: 749-751) explicitly rejects the view that "natural selection provides the missing causal link between what something does in a certain type of organism and its presence in that type of organism" (p. 750), but it is difficult to lay ones finger on exactly what view Cummins rejects and what exactly his arguments are. In the next section I will show that his discussion does not yield a tenable argument against my thesis that one might explain the presence of a certain kind of item in the lineage by appeal to the way in which past occurrences of that item performed their causal role.

Cummins's examples and his main thesis

Cummins starts his discussion of the view he rejects by considering two examples of attributions of causal roles:

> (a) The function of the contractile vacuole in protozoans is elimination of excess water from the organism
> (b) The function of the neurofibrils in the ciliates is coordination of the activity of the cilia (Cummins 1975: 749)

According to him such attributions have two distinct uses in biology. They serve (i) to explain specific capacities of individual organisms, and (ii) to explain how organisms that have items "of the sort in question" are able to survive.[13] For example, attribution (a) is used to explain (i) how certain organisms are able to get rid of water, accumulated by osmosis, and (ii) why protozoans that have contractile vacuoles do not explode in fresh water. Similarly attribution (b) is used to explain (i) how the activity of neurofibrils in ciliates is co-ordinated, and (ii) why ciliates incorporating neurofibrils are capable of fairly efficient locomotion. According to Cummins the impression that function attributions help to explain the presence of the "sort of

---

[13] Actually, Cummins's distinction lumps two contrasts together. One is the contrast between explaining a capacity of a certain individual and explaining a capacity of those individuals that are build in a certain way. The other is the contrast between explaining a specific capacity of an individual (c.q. those individuals that are build in a certain way) and explaining the ability of a certain individual (c.q. those individuals that are build in a certain way) to survive.

item” (p. 750) to which the function is attributed is due to a confused interpretation of the second use of function attributions:

> It might seem that natural selection provides the missing causal link between what something does in a certain type of organism and its presence in that type of organism. By performing their respective functions, the contractile vacuole and the neurofibrils help species incorporating them to survive, and thereby contribute to their own continued presence in organism of those species, and this seems to explain the presence of those structures in the organisms incorporating them (Cummins 1975: 750).

Cummins argues that the view stated in the quote above “involves a subtle yet fundamental misunderstanding of evolutionary theory” (p. 750).

Before considering Cummins’s arguments let me emphasize that I do not want to defend the view rejected by Cummins. Talk of “sort of item” and “type of organism” is too vague[14], and appeal to “survival of the species” is irrelevant. If this was the “subtle yet fundamental misunderstanding” Cummins talks about I would wholeheartedly agree with him. However, I do want to defend the view that the attribution of a causal role to an item sometimes helps to explain the evolution of that item and Cummins suggests that the misunderstanding is in this general idea, rather than in the wording he chooses. So, it seems that his arguments are also pertinent to my account. In that account the mechanism of natural selection provides the causal link between what an item did in certain organisms and its presence and character in the *lineage* of those organisms. Moreover, the link is provided not by the influence of the way in which a certain causal role is performed on the survival of the species (as Cummins suggests) but by the influence of the way in which a certain causal role is performed in a certain organism on the *relative fitness* of that *organism*.

Cummins’s argument centres around four claims: (1) the thesis that “the same” explanation might explain the presence of items that have quite different causal roles and survival values, (2) the thesis that the processes that lead to the incorporation of a certain item are insensitive to what that item does, (3) the thesis that the genetic plan of an organism is independent of the life chances of the organisms having that plan, and (4) the thesis that natural selection may destroy but not alter “genetic plans”. I shall discuss each of these claims and argue that they yield no argument against my account.

---

[14] It is not clear whether Cummins’s ‘sort of item’ consists of homologous items or of items with a similar structure. It is also not clear whether Cummins’s ‘type of organisms’ means ‘species’ or ‘the set of organisms that incorporate a certain item’ or ‘the set of organisms that incorporate a structural similar item’

Cummins's first claim

According to Cummins a clue to the mistake involved in the view that natural selection provides a link between "what something does in a certain type of organism and its presence in that type of organism" can be found in his example (a). In fresh water species the contractile vacuole has survival value because it enables the organism to get rid of the water that accumulates in the organism as the result of osmosis in organisms surrounded by a semi-permeable membrane. However, there are also marine species that have contractile vacuoles. These species do not have an excess-water problem but the reverse problem. Hence, the causal role and/or the survival value of contractile vacuoles are not the same in marine protozoans and in fresh-water species. However, according to Cummins, the explanation of the presence of contractile vacuoles in marine en fresh-water species is "almost certainly the same" (p 750). According to Cummins:

> This fact reminds us that the processes actually responsible for the occurrence of contractile vacuoles in protozoans are totally insensitive to what that [item][15] does (Cummins 1975: 750).

This first argument is far from clear. Cummins does not tell us what it is for two explanations to be "the same". He also does not state whether the explanations in question are concerned with the genesis of the contractile vacuole in the ontogeny (that is with the development of the item in the individual) or with the genesis of the contractile vacuole in the phylogeny (that is with the evolution of the item in the lineage). Moreover, he does not give bibliographical references to support his view.

It is easy to see that Cummins's first claim is irrelevant if it is concerned with explanations of the development of a contractile vacuole in the individual (whether it is true or not). Cummins might mean that the ontogeny of the contractile vacuole in marine protozoans is similar to the ontogeny of contractile vacuoles in protozoans that live in fresh-water. He might also mean that similar or homologous genes are involved in both ontogenies. Anyway, it is difficult to see how the claim that similar ontogenies may give rise to items with different causal roles and/or survival value would count against my view that the way in which a certain item performs its causal role may causally influence the character of that item in the lineage.

Let us, therefore, see what remains of Cummins's first claim if it is applied to the evolution of the item in the lineage. If we are to explain the evolution of contractile vacuoles the first thing to find out is whether or not contractile vacuoles in fresh water species are homologous to those in marine species. The second thing to find out is what the environmental conditions were that influenced the evolution of that/those item(s). Suppose that the contractile vacuoles evolved once (this might be what Cummins means if he says that the two explanations are the same) in,

---

[15] Cummins uses the term "structure" instead of "item".

say, a population of protozoans living in fresh water. The current species that incorporate contractile vacuoles descend from this ancestral population. If this scenario is true the causal role and survival value of the contractile vacuole in the current marine species is probably different from the causal role and survival value of that item in the ancestral population. Perhaps, in marine protozoans contractile vacuoles have no causal role at all; just as the human veriform appendix they are simply the remainders of a bygone past. Perhaps, the survival value of contractile vacuole is negative in the sense that it would be easier to survive for protozoans living in marine environments if they had no contractile vacuoles. This reminds us to the well-known fact that the causal role that helps to explain the evolution of a certain item need not be the causal role that item currently performs. However, it does nothing to undermine my view that the way in which a certain item performs its causal role could influence the presence and character of that item in the lineage.

Cummins's second claim

According to Cummins the "fact" (p 750) that "the two explanations are almost certainly the same" (p. 150) should remind us that "the processes actually responsible for the occurrence of contractile vacuoles in protozoans are totally insensitive to what that [item] does" (which is Cummins's second claim). This claim suffers from an ambiguity similar to the one in the first claim: it is not clear whether the phrase "the processes actually responsible for the contractile vacuoles in protozoans" refers to the genesis of vacuoles in the individual (that is to the ontogeny) or to the evolution of vacuoles in the lineage (that is to the phylogeny). It should be noted that the claim that the presence and structure of a certain item are insensitive to its causal role and survival value, applies to individuals at most. The processes responsible for the occurrence of a contractile vacuole in a certain protozoan individual are, perhaps, insensitive to what that vacuole does. But the processes at the population level that gave rise to the evolution of contractile vacuoles, were definitely not insensitive to what those items did. If those items were not advantageous to the individuals that had it, they probably would not have spread through the population.

Actually, even at the individual level it is in general not true that the processes responsible for the development of an item are totally insensitive to what that item does. There are many cases in which feedback mechanisms ensure that the character of an item is influenced by the manner in which that item performs its causal role. For example, it is well-known that the form of a bone is influenced by the forces operating on that bone and, hence by what that bone does. Another example is the influence of resistance on the ventricular walls: in mammals with a pathological high resistance in the pulmonary circuit, the right ventricular wall becomes thicker and resembles the left ventricle. What is true, of course, is that the causal role performed by a certain item in a certain organism and the survival value of that item for that organism do not influence the genotype of that organism. This is probably what Cummins means (see my discussion of Cummins's third claim in the next section).

In sum, Cummins's claim that "the processes actually responsible for the occurrence of contractile vacuoles in protozoans are totally insensitive to what that [item] does" can be interpreted in two ways. It might be seen as a claim about the emergence of contractile vacuoles in the course of the evolution or it might be seen as a claim about the emergence of contractile vacuoles in the course of the ontogeny. In the first case the claim is false. In the second case it is irrelevant to my account (which concerns evolution rather than ontogeny)

Cummins's third claim

The third claim can be found in the next paragraph of Cummins's paper. At this point Cummins observes that whether an organism incorporates a certain "sort of item", depends on whether that sort of item is specified by the "genetic 'plan'" (p 751) of that organism, not on its causal role or survival value. According to him this is clearly shown in case of mutations. If the genetic plan of an organism is changed by mutation the organism will incorporate the modified item, regardless of its causal role or survival value.

> The characteristics of organisms which determine their relative success or failure are determined by their genetic plan, and the characteristics of these plans are utterly independent of the relative success of organisms having them (Cummins 1975: 751).

Cummins's thesis that the genotype is independent of the relative success of the organisms having them is ambiguous, again.[16] If he means that the genotype of a certain organism is independent of the relative success of that organism he is certainly right. However, from this it does not follow that the presence of a certain genotype in the lineage is independent of the success of the organisms that have that genotype.[17] Natural selection cannot alter the genotype of

---

[16] In addition: the thesis that the characteristics of organisms that determine their relative fittness are determined by their genetic plan is, at best, too simple. One reason is that the relative fittness of an organism depends on the environment as well as on the constitution and behaviour of that organism. Another is that the constitution and behaviour of an organism in their turn depend on the environment as well as on the genetic constitution of that organism.

[17] A similar point is made by McClamrock (1993). McClamrock argues that

> His [Cummins's] arguments all depend on an illegitimate move from the observation that the causal (including functional) effects of a particular *instance* of a structure can't include the presence of that very structure (a trivial observation about the rejection of backwards causation) to the general (and incorrect) claim that a structure's having a certain function can't be a part of a causal account of the structure's presence (McClamrock 1993: 257, where Clamrock uses the term "structure" I would use "item").

an individual organism, but it can alter the genetic make up of the lineage. If a mutation influences the relative fitness of an organism its share in the population will change accordingly. For example, if a certain mutation is advantageous its share in the population increases and it might eventually become prevalent. If the increase in fitness was the result of the way in which a certain item performs its causal role that causal role helps to explain why the current organisms of a certain lineage incorporate an item different from its original item in its ancestors.

Hence, one again Cummins brings up a thesis that might be interpreted in two ways. If it is interpreted in one way (the genotype of a certain individual is independent of the relative success of that individual) it is true but irrelevant to my account (which is about the change of the genetic makeup in the lineage). If it is interpreted in the second way (the genetic makeup of the lineage is independent of the relative success of the individuals instantiating the genepool) it is relevant but false.

Cummins's fourth claim

This leads us to Cummins's fourth claim. Cummins acknowledges that the influence of a mutation on the fitness of an organism might influence the genetic makeup of a population. However, he maintains that this should not be seen as an alteration of what he calls "the genetic plan".

> If the alteration [introduced by a mutation in the "genetic plan"] is advantageous, the number of organisms inheriting that plan may increase, and, if it is disadvantageous, their number may decrease. But this has no effect whatever on the plan, and therefore no effect whatever on the occurrence of s' [the modified trait] in the organisms in question (Cummins 1975: 750).

This quote shows the now familiar ambiguity in the phrase "the organisms in question". As said before, Cummins is right that the success of a certain trait in a certain organism does not influence the genotype of that organism, but from this it does not follow that that success does not influence the occurrence of that genotype in the lineage.

In this part of his argument Cummins makes another error (in addition to the familiar juggling with the level of explanation). Cummins explains himself as follows:

> We could [...] think of natural selection as reacting on the *set* of plans generated by mutation by weeding out the bad plans: natural selection cannot alter a plan, but it can trim the set (Cummins 1975: 751, emphasis his).

Clamrock points out that all that is needed to see how the function of past instantiations of a certain item can influence current presence is an account of a mechanism. This account is provided by the theory of natural selection.

In this quote Cummins takes the metaphor of natural selection too literary. Natural selection primarily consists in heritable differences in fitness between the organisms of a population. These differences influence allele frequencies[18] in the genepool of that population. As a result of natural selection (that is as a result of individual heritable fitness differences) the relative frequency of a certain allele in the lineage may change:[19] it may decrease or increase.[20] A decrease may eventually result in the disappearance of the allele in question. As a result of an increase a certain allele may eventually become prevalent in the lineage. It is, therefore, wrong to view natural selection solely as a destructive force, as Cummins seems to do: natural selection decreases the frequencies of some alleles and increases the frequencies of others.

It is also wrong to think of the genepool of a population as a *set* of genotypes (as Cummins seems to do), because in doing so one ignores the *frequencies* of the different alleles. Since, changes in relative frequencies of alleles in the population are the primary effect of natural selection, it is a serious error to ignore them.

Cummins's contention that natural selection merely weeds the plans generated by mutation is, therefore, fundamentally mistaken. Both mutation and natural selection directly influence the frequencies of alleles in the population. Changes in the frequencies of alleles in the genepool of the population may lead to changes in the structure or activity of an item or behaviour (in the lineage) and, hence, to the appearance of new kinds of items or behaviours (that is to items or behaviours whose structure and activity differs from their predecessors). Natural selection (that is heritable differences in fitness of different variants in a population) is one of the processes that may effect changes in the genepool. Hence, natural selection is one of the causal factors in the appearance of new structures and behaviours (in the course of evolution).[21]

Cummins's conclusion

Cummins concludes his discussion of appeals to functions in evolutionary explanations as follows:

> Thus, we may be able to explain why a given plan is not a failure by appeal to the functions of the structures it specifies. [...] But this is not to explain why, e.g., contractile vacuoles occur in certain protozoan, it is to explain why the sort of protozoan incorporating contractile vacuoles occurs. Since we cannot

---

[18] Roughly speaking, alleles are different forms of the same gene.

[19] Natural selection may also result in stable polymorphism. I ignore this possibility in this paragraph, because I am concerned with explanations that appeal to directional selection, rather than with selection in general.

[20] Of course, the decrease of a certain allele at a certain locus is complementary to the increase of another allele at the same locus.

[21] As the results of plant and animal breeding testify, even in the absence of new mutation selection may lead to considerable morphological, physiological and behavioural changes.

> appeal to the relative success or failure of these organisms to explain why their genetic plan specifies contractile vacuoles, we cannot appeal to the relative success or failure of these organisms to explain why they incorporate contractile vacuoles (Cummins 1975: 751).

This passage shows Cummins caught in his own words. For, assume that "the sort of protozoan incorporating contractile vacuoles" refers to the subclass of protozoans that have contractile vacuoles, then to explain "why the sort of protozoan incorporating contractile vacuoles occurs" (why there are protozoic individuals that have contractile vacuoles) is the same as to explain "why contractile vacuoles occur in certain protozoan". On the other hand assume that "the sort of protozoan incorporating contractile vacuoles" refers to the lineage(s) of protozoans having contractile vacuoles, then Cummins's remark that to explain "why a given plan is not a failure by appeal to the functions of the structures it specifies" is to explain "why the sort of protozoan incorporating contractile vacuoles occurs" (why there evolved lineages of protozoans with contractile vacuoles) is false. On the contrary, it was one of Darwin's important insights that to explain why organisms that are built in a certain way are viable, is not sufficient to explain why such organisms evolved.

5.5.3 Conclusion

I have argued that appeals to causal roles (function$_2$ attributions) in explanations of the evolution of the item or behaviour to which the causal role is attributed are acceptable on a causal account of explanation. The way in which a certain item or behaviour in a certain individual performs its causal role influences the fitness of that organism. Differences in the way in which homologous items or behaviours in different individuals in a population perform their causal role, may lead to differences in fitness between those organisms. These differences may lead to a subsequent change of the character of those items or behaviours in the lineage. Thus, natural selection provides the mechanism by which the causal role of an item or behaviour influences the presence and character of that item in the course of evolution.

Cummins seems to disagree with this view and I have considered his arguments. I have shown that these arguments suffer from a number of obscurities, such as vague talk of "sorts of items" and "types of organisms", drivel about "survival of the species", unsubstantiated and vague claims about two explanations being "almost certainly the same", a recurrent failure to handle the distinction and the interaction between population level and individual level processes correctly, and a mistaken view of natural selection as trimming a set of genetic plans. A charitable reading of his argument may yield two tenable thesis, namely (1) the causal role that influenced the evolution of an item or behaviour might not be the causal role that item or behaviour currently has, and (2) the way in which a certain item or behaviour in a certain organism influences the survival of that organism does not influence the genotype of that organism. Both these theses are irrelevant. I, therefore, draw the conclusion that Cummins's argument gives us no reason to doubt the thesis that natural selection provides the mechanism by which the causal role of an item or behaviour influences the presence and character of that item or behaviour.

5.6 Discussion of some criticisms of Cummins's account

My main interest is in an account of functional explanation. In this section I deal with two issues in regard to Cummins's notion of function. First, I discuss Amundson & Lauder's (1994) defence of Cummins's account of function. Second, I discuss the idea that Cummins's analysis provides us with a necessary but not with a sufficient condition to attribute a causal role to an item.

5.6.1 Amundson & Lauder (1994)

In a paper titled "Function without purpose" (1994) the philosopher Ron Amundson and the morphologist George Lauder co-operate in a defence of Cummins's account of function. More specifically, they argue (i) that Cummins's account of function closely matches the concepts of function used in functional morphology, (ii) that the philosophical criticisms of Cummins's concept of function put forward by Millikan (1989b), Neander (1991a) and Sober (1993) do not hold water, (iii) that functions as causal roles can not be eliminated from functional biology in favour of functions as selected effects, and (iv) that functions in Cummins's sense are essential to certain important research programs in evolutionary biology.

Unfortunately Amundson & Lauder fail to distinguish between function$_1$ (activity) and function$_2$ (causal role). In this section I discuss their argument ad (i) and ad (iv). Ad (i) I argue that Amundson and Lauder are right that Cummins's account of function closely matches many uses of the term 'function' in functional morphology (see section 5.3) but they do not show that. Instead, they point out that the notion of function as activity (function$_1$) is important in functional morphology. Ad (iv) I argue that Amundson and Lauder point to important uses of both the notions of function as activity (function$_1$) and function as causal role (function$_2$) in evolutionary biology. Strange enough they do not mention the appeal to causal roles in selection explanations (see section 5.5 above). I return to their argument ad (ii) the next section (5.6.2) and to their argument ad. (iii) in section 7.3.1.

Amundson & Lauder's main argument that Cummins's account matches the concepts of function used in functional morphology rests on Walter Bock and Gerd von Wahlert's (1965) account of the vocabulary of functional morphology.[22] Amundson & Lauder point out that

---

[22] I have discussed Bock and Von Wahlert's paper in section 3.1.

according to Bock & Von Wahlert's the form and the function are both at the "methodological base level" (p. 449) of morphological research: form and function are observable qualities which are not inferred from an evolutionary narrative, but, on the contrary, serve as the starting point to address questions about survival value and evolutionary history. Without such a base research in survival value and evolutionary history could not get off the ground. According to Amundson & Lauder Cummins's concept of function is similar to that of Bock & Von Wahlert's because both define 'function' without invoking evolutionary history or survival. And, indeed, Bock & Von Wahlert and Cummins agree on this point. However, there is an important difference between these accounts. Amundson & Lauder rightly point out that according to Bock & Von Wahlert the functions of an item are observable qualities and function attributions state what an item does. However, in Cummins's account functions are not observable properties, but capacities singled out by their role in capacity explanations, that is by the way in which a certain item is used.[23] Amundson and Lauder are aware of this difference where they observe that Bock & Von Wahlert's account is more radical than Cummins's.

> In one way, Bock and von Wahlert's concept of function is even more radical than Cummins's. Cummins assigns functions only to those capacities of components which are actually invoked in a [capacity explanation],[24] those which are believed to contribute to a higher level capacity to be analyzed. Bock and von Wahlert include *all possible* capacities (causal powers) of the [item],[25] given its current form. (Amundson & Lauder 1994: 450).

However, they set this issue aside with the following words:

> Apart from the issue of unutilized functions, Cummins's concept of function matches the anatomists' (Amundson & Lauder 1994: 450).

Actually, this issue is crucial both from the point of view of Bock & Von Wahlert and from that of Cummins. As I discussed in section 3.1, one of the central issues in Bock & Von Wahlert's account is the distinction between "function" and "biological role", words they use to refer to the distinction between what an item does and what it is used for. The distinction is crucial from Cummins's point of view because according to Cummins function attributions owe their meaning to capacity explanations. There would be no reason to call an activity of an item a function if that activity had no role in a capacity explanation. Hence, Amundson & Lauder do not show what they purport to show, namely that functions in Cummins's sense play an important role in functional biology. My examples in section 5.3 show that they are nevertheless

---

[23] In other words: a function in Cummins's sense is not a property but a position in a system that performs a certain task.

[24] Amundson & Lauder say "functional explanation".

[25] Amundson & Lauder, following Bock and Von Wahlert, say "feature"

right. What they do show is that the notion of function as potential activity (function$_1$) plays an important role in functional biology. This point is well worth making.

In section 7 of their paper, Amundson & Lauder argue that Cummins's concept of function is essential to certain important research programs in evolutionary morphology. These research programs are not concerned with selection. Amundson & Lauder list three kinds of research programs in which Cummins's concept of function is essential: (i) attempts to reconstruct phylogenetic trees, (ii) explanations that show how certain important evolutionary changes were made possible by so-called key transformations, (iii) explanations of why certain evolutionary pathways were not taken by appeal to the constraints put on evolution by the causal role of the items involved. The first kind of research program is concerned with the reconstruction of the evolutionary history, rather than with its explanation. Amundson & Lauder point out that functional characteristics can be used to reconstruct phylogenetic trees in exactly the same way as morphological characteristics. The two other kinds of research programs seek to explain evolutionary history. An example of a key transformation is the duplication or repetition of parts of the body (i.e. segmentation). Segmentation appears to be a vehicle for the generation of evolutionary diversity because it facilitates independent specialization of the components (Lauder & Liem 1989). Another example (not mentioned by Amundson & Lauder) is the development of a circulatory system. This transformation facilitates the development of larger organisms (Krogh 1941). In regard to the third kind of research program, Amundson & Lauder point out that in generally many items work together to perform a certain activity, for example, jaw movement. In such case minor changes in form might have a deleterious effect on the performance of that role.

Amundson & Lauder's examples of the second and third kind of research programs show that the study of functional interdependencies plays an important role in explanations in evolutionary biology that do not centre around selection. This underscores my thesis that the central notion of function in functional biology is the notion of function as causal role. Amundson & Lauder give the impression that the concept of function as selected effect does play an important role in selection explanations in ethology. I contend that even in selection explanations the relevant concept of function is function as causal role (see especially section 5.5 and 7.2).

In sum: Amundson & Lauder give interesting examples of the use of the concept of function as activity (function$_1$) in both functional biology and evolutionary biology, as well as of the importance of the concept of function as causal role (function$_2$) and of the concept of functional interdepencies in explanations in evolutionary biology.

5.6.2 Is Cummins's analysis in need of suppletion?

In section 5.3.2 I argued that Cummins's idea that function attributions are singled out by their role in capacity explanations applies to function$_2$ attributions (attributions of causal roles).

In the philosophical literature surface a number of counterexamples that suggest that not all capacities singled out by a capacity explanation should count as functions and, hence, that Cummins's analysis should be supplemented with another condition. Candidates for such an additional condition are the condition that the exercise of the causal role has positive survival value and the condition that the exercise of the causal role has contributed to the evolution of the item to which the function is attributed. The counterexamples fall into two kinds. The first kind of examples consists of examples in which Cummins's analysis would single out a certain capacity of a part of an organism as a function, whereas one normally would not call the exercise of that capacity a function of that part. The second kind consists of examples in which Cummins's analysis would allow us to attribute functions to parts of systems of which one normally does not attribute functions to their parts.

A counterexample of the first kind is brought up by Neander (1991a: 181) . Neander contends that the pressure exercised by a tumour on a certain artery in the brain helps to explain an organism's complexily achieved capacity to die of cancer. Yet, this causal role does not count as its function. According to Neander tumours don't have functions at all. Sober (1993: 86) advances a similar counterexample. He objects that on Cummins's account one of the functions of the heart is to weigh what it does because this contributes to the organism's capacity to tip the scales on a certain number of pounds.

Counterexamples of the second kind are furnished by Millikan and Neander. Millikan (1989b: 294) states that on Cummins's account in the context of the water-cycle system "it is arguably the function of clouds to make rain with which to fill the streams and rivers", for this helps to explain "how moisture is maintained in the soil so that vegetation can grow". According to Neander (1991a: 181) Cummins's analysis bestows functions to plate movements, for they help to explain earthquakes.

None of Cummins's critics document their examples, nor does any of them details the capacity explanation that would justify the counterintuitive function attribution in question.

As Amundson and Lauder (1994: 452) have pointed out some of these purported counterexamples must be rejected outright: for on Cummins's criteria (see section 5.2.2) there is no reason to apply the strategy of functional analysis to the systems in question. The explaining capacities would not be simpler than and/or would not differ in kind from the capacities to be explained and the organization of the explaining capacities would not be very complex. Amundson and Lauder take Neander's contention that Cummins's account allows us to attribute functions to plate movements as their main target. They answer that such attributions are not allowed because plate movements are not simpler than earthquakes and because although the geological structures that bring about earthquakes are complex, the structure of the capacities that would explain the earthquake would be simple.

Sober’s alleged counterexample (“a function of the heart is to weigh what it does”) is another clear case that should be rejected on the ground that there is no reason to suppose that the relevant capacity (i.e. the capacity of the heart to weigh what it does) is singled out by a suitable capacity explanation. An explanation of the capacity of a certain organism to tip the scales on a certain number of pounds would start with the observation that scales measure weight. So to explain the capacity of a certain organism to tip the scales on a certain number of pounds one should explain the weight of that organism. One may explain the weight of an organism in terms of the weights of its parts and the fact that weights are additive. This would be an analytical explanation but not a capacity explanation. The analysis on which the explanation is based is componential rather than functional.[26] After all, the property of the parts that helps to explain the weight of the organism is the same for all parts: that part’s weight. On Cummins’s account a mere componential analysis does not justify the attribution of functions to the parts. Therefore, the explanation of the capacity of an organism to tip the scales on a certain number of pounds by appeal to the weight of its heart does not justify to attribute to the heart the function to weigh what it does.

Millikan’s purported counterexample (“it is the function of clouds to make rain with which to fill the streams and rivers”) is worthless as it stands. According to Millikan, the explanation that justifies the function attribution explains “how moisture is maintained in the soil so that vegetation can grow”. It is not at all clear what capacity the purported capacity explanation would explain (is ‘so that’ part of the description of a capacity?) and in which system that capacity is instantiated, still less how that explanation would proceed (what exactly are the explaining parts and what the explaining capacities?). Millikan’s words suggest that the capacity to be explained Millikan has in mind is something like the capacity of the soil to supply the vegetation it has with enough water. But in that case it is not clear how such an explanation can attribute functions to clouds, which are not part of the soil. This consideration suggests that the capacity to be explained is to be attributed to the atmosphere. It is not difficult to imagine a capacity explanation of the capacity of the atmosphere to sustain life on land that appeals to the capacity of clouds to produces rain. Such an explanation would analyse the ability to sustain life on land into a number of subcapacities such as the ability to maintain a certain temperature, the ability to maintain a water cycle, the ability to shed ultra violet radiation and so on. The ability of the atmosphere to maintain a water cycle in its turn is analysed into the ability of clouds to produce rain, of streams to transport water to the sea, of soil to accumulate ground water, of water surfaces to evaporate water, of organisms to transpirate and so on. If someone comes up with such

---

[26] The capacity of the scales to measure weight can be suitable explained by means of a capacity explanation. However, such an explanation would bestow functions to the parts of the scales, but not to the weights it measures.

an explanation no one would have hesitations to speek of functions. Indeed, in papers discussing the earth's economy it is quite common to talk of the functions of e.g. soil, sea and clouds in e.g. the water cycle (see for example Keeton & Gould (1993: 1154-1169) and Rimmer (1998). It seems that Millikan's counterexample turns into an example if the relevant capacity explanation is properly worked out.

Neander's example of cancer tumours is the most interesting one. However, cancer is a difficult subject and this book seems not the proper place to discuss it. As long as Neander herself has not substantiated her claim I see no reason to accept it.

I draw the conclusion that none of the criticisms mentioned discussed here gives us reason to doublt Cummins's view that attributions of causal roles are singled out by their role in capacity explanations.

5.7 Conclusion

Cummins distinguishes two kinds of explanation: explanation by subsumption and analytical explanations (see section 5.2.2 above). Explanations by subsumption are used to explain changes as the result of preceding changes. Analytical explanations are used to explain the properties of a certain system in terms of the properties of the components of that system and their mode of organization. According to Cummins, functional explanations are a subspecies of the category of analytical explanations. Functional explanations explain a dispositional property (capacity) of a complex system. In addition to the componential analysis, they provide a functional analysis of that capacity: they analyze the capacity to be explained into a series of subcapacities (instantiated in the components) that add up to the capacity to be explained.

Cummins's account enables us to understand the structure and the explanatory force of capacity explanations. Such explanations show us how the ability of a certain item or behaviour to perform a certain complex task results from the capacity of the parts of that item or behaviour to perform a series of more simple tasks which add up to the complex ability. This procedure is repeated until a level is reached at which the simple capacities might be explained by the physical and chemical of the parts and the way in which they are organized. Such explanations are explanatory because they show us how a complex capacity is brought about by the co-ordinated activity of simple building blocks.

Attributions of causal role owe their scientific meaning to this strategy of explanation. They tell us what tasks a certain item or behaviour has. In that way they situate an item or behavioural pattern in the organism's organization.

Cummins restricts his account of 'function' to the use of attributions of causal role in capacity explanations . I show that attributions of causal role have other uses too. If the place of an item or behaviour in the animal's organization is known this knowledge can be used to explain the presence and character of the item to which the function is attributed in two ways, namely in design explanations and in selection explanations.

Design explanations fall outside the scope of Cummins's categories. They explain the properties of an item or the behaviour of an organism. They do so by relating those properties to other properties of the organism in question and to the state of the environment in which it lives, not by analyzing that item into components. This means that Cummins's account does not help us to understand design explanations.

Selection explanations do explain changes and the appeal to causal roles in this kind of explanations is explanatory on a causal account: selection for a more efficient performance is one of the processes that may explain how a certain kind of items is brought about in the course of evolution.

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 uncontroversial. 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.

Chapter 7: The etiological theory

7.1 Introduction

In chapter 4 I argued that Hempel's and Nagel's classical accounts of design explanations in terms of the inferential theory of explanation, fail. Generations of philosophers have talked about these attempts as if they were "the received view" (the last representative of this tradition is probably Sandra Mitchell (1989: 214)). The expression 'received view' is probably used merely as a proper name, rather than as a definite description. In this chapter I discuss what might be called the received view of our time: the "etiological account". On the etiological account function attributions in biology inform us about the evolutionary history of the item to which the function is attributed. More specifically, function attributions identify the effects for which a certain trait was selected in the past. For example, on the etiological account hearts are said to have the function of propagating the blood, if and only if propagating the blood is what hearts did that caused them to be favoured by natural selection.

The etiological account as it is defined here was originally proposed by Karen Neander (1980, 1983, 1991a, 1991b), and, independently, by Ruth Millikan (1984, 1989b, 1989a, 1993a). Other proponents of etiological theories are Mitchell (1989), Brandon (1990), Griffiths (1993), Sober (1993), and Godfrey-Smith (1994). Forerunners of the etiological theory were proposed by Ayala (1970) and Wright (1973, 1976). According to the latter theories, functions are determined partly, but not wholly by a selection history. Millikan (1993a: 33) claims that she arrived at her theory without knowing Wright's work and she tends to play down the similarities between an etiological account and Wright's theory of function. The other authors generally acknowledge Wright's influence on their work.

The etiological theory of function is first and foremost concerned with the definition of the term 'function'. The role of function attributions in explanations receives much less attention. However, the etiological account has a very attractive feature for philosophers who favour a causal theory of explanation: the etiological account promises to reconcile the idea that function attributions concern consequences with the idea that function attributions are explanatory on a causal theory of explanation. The designer of the causal theory, Wesley Salmon, was among the first to appreciate this connection between the causal theory of explanation and the etiological account of function (Salmon 1989: 111-116). The point has been made most forcefully by Neander (1991b) and by Mitchell (1993). At first sight function attributions seem to be both explanatory and concerned with consequences. However, on the causal theory of explanation, explanations are explanatory to the extent that they detail the mechanisms by which the phenomenon to be explained was produced. But how could an appeal to consequences causally explain the presence of an item, since those consequences occur only after the production of that item? The answer of the etiological theory is as ingenious as it is simple: the functions of a certain item in a certain individual are a subset not of the present consequences of that item but rather of the past consequences of that item in ancestral organisms. Past consequences, of course, can be causally effective and the etiological account identifies the functions of an item with those past consequences that *were*, as a matter of fact, causally effective in the evolution of the item to which the function is attributed.[1]

Both Neander and Millikan claim that function attributions have an indispensable role in categorizing items and behaviours and that among the current theories of function the etiological theory is the only one that accounts for this role of function attributions. According to them biological categories such as ‘heart’, ‘kidney’ and ‘mating display’ collect items together on the basis of their function. It is obvious that a certain biological category may include normal members that do not actually perform their function (e.g. many attempts to escape fail and most sperm cells will never fertilize an egg cell) as well as defective members that are not even able to perform their function (a malformed heart is still a heart). The etiological theory accounts for this feature by defining the function of an item or behaviour not on the ground of the actual properties and propensities of that item or behaviour but in terms of its history. On the etiological theory a certain sperm cell is a sperm cell because fertilizing eggs is what the precursors of that cell did that accounts for the existence of the sperm cell in question. Similarly, propagating the blood is the hallmark of hearts and a malformed heart is yet a heart because propagating the blood is what precursors of that malformed heart did that accounts for the existence of that malformed heart.

In this chapter I am concerned with the question to what extent the etiological theory accounts for the use of the different kinds of function attributions in functional biology. I focus on Millikan’s account because this is currently the most elaborated etiological account of function. As she has emphasized recently (Millikan 1993a: 31), her definition of what she calls “proper functions” is not meant to capture the biologist’s usage. Millikan is primarily interested in solving certain problems in the philosophy of language (Millikan 1984) and in the philosophy of mind (Millikan 1993b). Her definition of ‘proper function’ is meant as a stipulated, technical term, the use of which does not depend on the extent to which it captures the use of

---

[1] Neander (1991b: 462) rightly points out that this is true only of biological functions. Functions of human artefacts are effects of which the designer or user believes or hopes that they will occur and for which the item is selected. Such effects need not have occured in the past. Since I aim to account of biological functions only I shall ignore this point.

function attributions in biology. Nevertheless there are three good reasons for considering the question whether or not her theory applies to biological function.

The reason for considering Millikan's theory is obvious: Millikan has offered a powerful general account of function. This alone is sufficient reason to consider its application to biological function.

The second reason is that many philosophers of biology have arrived at accounts similar to that of Millikan (Mitchell 1989, Brandon 1990, Neander 1991b, Neander 1991a, Griffiths 1993, Sober 1993). All these authors claim that when biologists attribute a function to an organ, process, behavioural pattern, characteristic or whatever they make a claim about the evolutionary history of those items or traits.

The third reason is that Millikan herself repeatedly suggests that her theory applies to almost all function attributions in biology. Some of her favourite examples of function attributions are 'the heart has the function to pump the blood around', 'sperm cells have the function to fertilize an egg cell' and 'the colour of the chameleon's skin has the function to camouflage its bearer'. How could such examples make sense if Millikan does not think that her theory applies to these examples? Indeed in her "In Defense of Proper Functions" (1989b: 293) Millikan explicitly claims that her theory explains all ordinary uses of the terms 'function' and 'purpose', including their use in biology:

> The definition of "proper function" is intended as a theoretical definition of function or purpose. It is an attempt to describe a unitary phenomenon that lies behind all sorts of cases in which we ascribe purposes of functions to things, which phenomenon normally *accounts for* the existence of the various analogies upon which applications of the notion "purpose" of "function" customarily rest. My claim is that actual body organs and systems, actual actions and purposive behaviors, artifacts, words and grammatical forms, and many customs, etc., all have proper functions and that these proper functions correspond to their functions or purposes ordinarily so called. Further, it is *because* each of these has a proper function or set of proper functions that it has whatever marks we tend to go by in *claiming* that it has functions, a purpose, or purposes (Millikan 1989b: 293).

In the same year she publishes a paper in *Biology and Philosophy* (Millikan 1989a) in which she distinguishes two different kinds of function: "Cummins-style functions" and "selected functions". The latter ones are also called "proper functions". These kinds of functions correspond to the kinds of functions I call 'function as causal role' (function_2) respectively 'function as selected effect' (function_4). Associated with these two kinds of function Millikan distinguishes two kinds of functional explanation: "functional explanations in Cummins sense" and "functional explanations that make reference to natural selection". These two kinds of functional explanation correspond to the kind of explanations I call 'capacity explanations' respectively 'selection explanations'. Millikan maintains that the use of Cummins-style functions is restricted to functional explanations in Cummins sense, whereas selected functions may “enter into” functional explanations that make reference to natural selection. Although Millikan seems to acknowledge that both kinds of function have their use in biology she still maintains that selected functions are “the only kind of function that is well defined” (Millikan 1989a: 174). Indeed, four years later she argues that Cummins’s definition of function does not suffice to define biological function (Millikan 1993a: 33, 35-39).

I will argue that the etiological theory is not relevant to understand reasoning about functions in functional biology.

7.2 Proper functions and selection

7.2.1 The etiological account of function

In the first two chapters of her *Language, Thought and Other Biological Categories* (1984) Millikan develops a recursive definition of what she calls “proper function”. This definition is a general definition applicable to such diverse things as organs, behaviours, artefacts, customs, parts of organizations, words and grammatical forms. Although Millikan’s examples include functions of organs of organisms such as hearts and kidneys, her intended domain of application is the philosophy of mind and language. As I am interested only in the application of the theory to biology, I restrict this summary to those definitions that allow one to attribute functions to parts and behaviours of living organisms.

On Millikan’s theory functions are attributed to things that have been reproduced repeatedly, that have been selected to be reproduced, and that were selected because of what they did. The proper function of such a device is to do what its ancestors did that explains why those ancestors were selected for reproduction. That is, to do the things that explain why the device exists, why it exists in the place where it exists, or why it is as it is:

> Putting things very roughly, for an item *A* to have a function *F* as a “[direct][2] proper function”, it is necessary (and close to sufficient) that [...] *A* originated as a “reproduction” (for example, as a copy, or a copy of a copy) of some prior item or items that, *due* in part to possession of the properties reproduced, have actually performed *F* in the past, and *A* exists because (causally historically because) of these performances (Millikan 1989b: 288/9).

---

[2] Actually, Millikan gives two disjunctive conditions. The first condition is satified by "direct proper functions", the second by "derived proper functions" (functions derived from the proper functions of the devices that produce them). I have quoted only the first condition. Derived proper functions are not of interest here.

The definition of the proper function of the parts and behaviours of organisms is complicated by the fact that the parts and organs of a certain organism are not copied from the parts and organs of the parents of that organism. It is the genes that help to produce those items which are copied, not the items themselves (see fig. 7.1). This means that, to attribute proper functions to the organs of organisms the notion of reproduction must include more than copying. Millikan needs a notion of reproduction that allows, for example, my lung to be a descendent of the lungs of my parents.

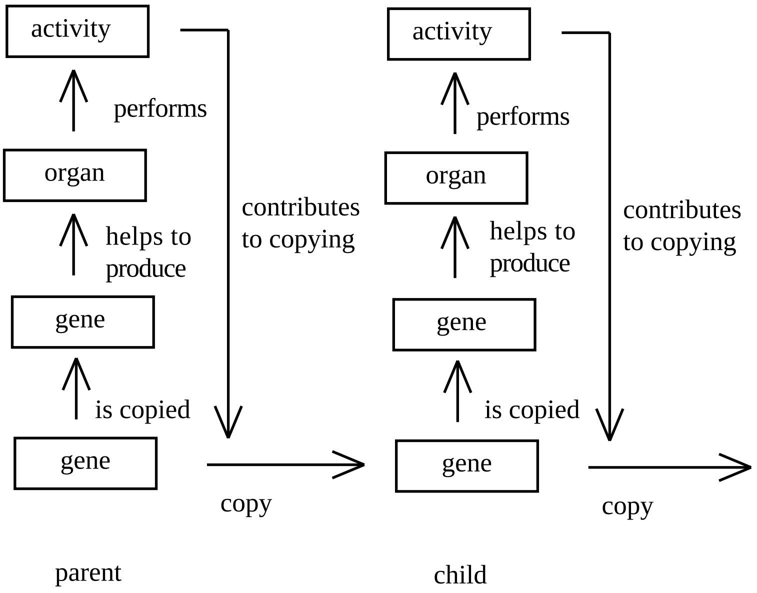

Fig 7.1: reproduction in organisms

Millikan solves this problem by means of a recursive definition. Proper functions are defined with respect to entities that have been reproduced. The class of entities that have been reproduced includes (i) entities that have been copied, and (2) entities that have been produced with help of entities that have been reproduced and that have the proper function to help to produce those entities. Genes belong to the first category of entities that have been reproduced, organs to the second.

Now for the details. The recursion starts with entities, such as genes, which are copied. The hallmark of copying is that the copies are similar to the original. The copies of an entity which is copied establish a “first-order reproductively established family”:

> Any set of entities having the same or similar reproductively established characters derived by repetitive reproductions from the same character of the same model or models form *a first-order reproductively established family* (Millikan 1984: 23).

According to this definition the subsequent copies of a gene constitute a reproductively established family, provided that those copies are sufficiently similar to each other.

Next, Millikan uses the definition of proper function to attribute functions to entities that originate from copying. In order to have proper functions the entities must not only originate from copying (or from other kinds of reproduction), there also must have been a kind of selection process. That is, there must have been competitors with a different character which were not reproduced or were less extensively reproduced because due to their different character they performed a certain activity less well. According to Millikan's definition of proper function an entity that originated as a reproduction of a reproduction (...) of another entity has an activity *F* as a proper function if (i) some of its ancestors competed to be reproduced with entities of a different character, (ii) the proportion of entities that did *F* was higher among the entities with the ancestor's character than among the entities with the different character, and (iii) those ancestral entities were selected for reproduction over their competitors because they did *F*. The character due to which ancestral entities performed *F* and, hence, due to which they were selected, is called the 'reproductively established' or 'Normal' character of that type of entity.

> Where *m* is a member of a reproductively established family *R* and *R* has the reproductively established or Normal character *C*, *m* has the function *F* as a direct proper function iff:
>
> (1) Certain ancestors of *m* performed *F*
>
> (2) In part because there existed a direct causal connection between having the character *C* and performance of the function *F* in the case of these ancestors of m, *C* correlated positively with *F* over a certain set of items *S* which included these ancestors and other things not having *C*.
>
> (3) One among the legitimate explanations that can be given of the fact that *m* exists makes reference to the fact that *C* correlated positively with *F* over *S*, either indirectly causing reproduction of *m* or explaining why *R* was proliferated and hence why *m* exists (Millikan 1984: 28)

Millikan does not give an example of an attribution of a function to a member of a first-order reproductively established family but one might think of *m* as a gene of a certain organism (let's call that organism *i*) and of *C* as the order of the nucleotides of that gene. *F* might be the production of a certain peptide, say haemoglobin. *R* consists of all the genes in *i* and in related organism which are both homologous with and sufficiently similar to *m*. According to the above definition gene *m* of *i* has the function to produce haemoglobin if and only if (1) some ancestors of *m* produced haemoglobin, (2) beside the ancestors of *m* with nucleotide sequence *C* there existed variant genes with a different sequence; because the ancestors had nucleotide sequence *C* they produced haemoglobin whereas the variants did not produce haemoglobin (or less haemoglobin, or a less efficient kind of haemoglobin), (3) the existence of *m* might be explained by appeal to the fact that (due to the fact that the order of their nucleotides was *C*) the ancestors of *m* produced haemoglobin whereas the others did not (or produced less haemoglobin, or less efficient haemoglobin). Suppose there is reason to assume that from time to time there occur mutants of the gene for haemoglobin that produce a less effective kind of haemoglobin (think of genetically induced anaemia) and that the fitness of those variants is lower than that of the normal kind of haemoglobin. Then we might also assume that some ancestors of *m* were in competition with such mutants and we might explain the existence of *m* by appeal to the differences between the normal variant and the mutant.[3] Hence, we might attribute the proper function to produce haemoglobin to *m*.

This shows that the above definitions of first-order reproductively established family and of proper function allow us to attribute proper functions to genes. In order to be able to attribute functions to other items such as organs, Millikan introduces the notion of a “higher-order reproductively established family”. This definition uses the notions of proper function and of Normal explanation. A Normal explanation tells us how ancestral items performed the activity for which they were selected, for example a Normal explanation might explain how past copies of a gene were used to produce haemoglobin. The conditions that enabled ancestral items to perform the activity that resulted in selection of that item are called the ‘Normal conditions’.

> Any set of similar items produced by members of the same reproductively established family when it is a direct proper function of the family to produce such items and these are all produced in accordance with Normal explanations, form *a higher-order reproductively established family* (Millikan 1984: 24)

On this definition the parts and organs of an organism form a reproductively established family. As I showed above, on the basis of the definitions of first-order reproductively established family and of proper function one may attribute the proper function to produce a certain molecule (such as haemoglobin) to a certain gene. Because that gene is a member of a reproductively established family and because the members of that family have the proper function to produce haemoglobin, haemoglobin is a member of a higher-order reproductively established

---

[3] Note that this is my reconstruction of what Millikan would say of this example, not my own opinion. Part (3) of Millikan’s definition of proper function requires that the existence of *m* can be explained by appeal to selection. For that reason, I assume that Millikan would claim that the selection explanation in my example explains why *m* exists. I would not agree with her. The selection explanation in my example explains how the character of the members of *R* was maintained in the evolutionary history. This explains why the members of *R* have the character they have and, hence, why *m* has the character is has, but not why *m* exits. Indeed, I fail to understand what it would mean to explain why *m* exists by appeal to selection. Because I feel awkward in saying that the selection explanation in my example explains the existence of *m* and because I would agree that *m* has the proper function to produce haemoglobin I think that part (3) of Millikan’s definition of proper function is defective. This defect can be repaired easily by replacing the requirement that one may explain why *m* exists by appeal to selection with the requirement that one may explain why members of *R* have the character they have by appeal to selection.

family. Similarly, hearts form a higher-order reproductively established family because there are several genes that have the proper function to (help to) produce hearts.

As the definition of proper function applies to all kinds of reproductively established families, it is now possible to assign functions to organs and behaviours. Because my heart is a member of a reproductively established family, it has the proper function to pump blood if pumping blood is what its ancestors did that account for its presence in my body. In that case *m* refers to my heart, *C* to those properties of hearts that enable it to pump blood, *F* is propagating the blood and to *m* belong all hearts which are both homologous and sufficiently similar to each other.

As I said, Millikan's theory is meant as a general theory which allows one to attribute proper functions to a lot of entities, provided that these entities are reproduced and that their structure or existence can be explained by appeal to selection. When applied to the living world the relevant selection process is natural selection. The term 'natural selection' refers to the situation that (i) there is variation among the individuals of a population in some trait, (ii) the life chances of these variants vary as a result of that variation (the different variants differ in fitness due to the different possession of that trait), and (iii) the possession of that trait is inheritable. In short, natural selection consists in heritable fitness differences. Natural selection results in (i) differences in the frequency distribution of the characteristic among age classes, and (ii) a change in relative frequency of that characteristic among generations (if the population is not at equilibrium). Note that these effects might also result from other sources. Differences in trait frequency distribution among age classes might be the result of ontogenetic development (the fact that the frequency of baldness increases with the age is for example not an effect of selection), changes in trait frequency among generations might be the result of direct environmental effect or of genetic drift.

Whereas Millikan pretends indifference to the applicability of the etiological theory to biology, Neander's theory is explicitly meant as an analysis of the biologist's usage. Neander defines proper functions directly in terms of natural selection. Neander defended this theory in a paper she presented at the AAP conference in 1980 (Neander 1980) . That paper circulated widely among philosophers of science with an interested in biology but remained unpublished. Neander elaborated on the etiological theory in her PhD thesis (Neander 1983). Her first published defence of the etiological theory appeared in 1991 in *Philosophy of Science* under the title "Function as Selected Effects":

> This paper defends the etiological theory of proper functions, according to which, roughly speaking, biological proper functions are effects for which trait were selected by natural selection. (Neander 1991a: 196).

More specifically:

> It is the/a proper function of an item (*X*) of an organism (*O*) to do that which items of *X*'s type did to contribute to the inclusive fitness of *O*'s ancestors, and which caused the genotype, of which *X* is the phenotypic expression, to be selected by natural selection. According to this theory, for instance, hearts have their proper function of pumping blood, because pumping blood is what hearts did that caused them to be favored by natural selection (Neander 1991a: 174).

For example

> it is the function of your opposable thumb to assist in grasping objects, because it is this which opposable thumbs contributed to the inclusive fitness of your ancestors, an which caused the underlying genotype, of which opposable thumbs are the phenotypic expression, to increase proportionally in the gene pool. In brief, grasping objects was what the trait was selected for, and that is why it is the function of your thumb to help you to grasp objects (Neander 1991a: 174).

7.2.2 Selection explanations

It will be clear from the above account that on the etiological theory there is a tight connection between attributions of proper functions and explanations that appeal to natural selection. In this section I take a closer look at such explanations. I distinguish two kinds of selection explanation: evolutionary selection explanations and equilibrium selection explanations. Evolutionary selection explanations explain the presence or character of a certain item or behaviour by telling how and why that item or behaviour was modified by natural selection in the course of evolution (see section 2.3.2). Equilibrium selection explanations explain why a certain variant is maintained by natural selection in the population at a certain frequency.

In evolutionary selection explanations the present character of an item is viewed as the result of a series of changes of an ancestral item. Each of these changes is explained as the result of natural selection, which is a process that operates at the population level (this is what distinguishes evolutionary explanations from physiological, developmental and design explanations, all these explanations are concerned with processes, mechanisms and/or relations at the individual level)

A well-known example of an explanation of an evolutionary change by appeal to natural selection is Kettlewell's (1959) explanation of the increase of melanic (black) variants of various species of moths in the industrial areas of nineteenth century England. In the area around Manchester the first black variants of the peppered moth (*Biston betularia*) were caught in 1848; by 1895 this variant took up 95% of the total population in that area. Kettlewell explains this increase as follows. The various species of moths rest during the day on the exposed parts of trees. The resting moths form a favourite dish for many insect eating birds. Before the rise of industrialization the trees were covered with light-coloured lichens. Against this background the light forms of the moth were difficult to see. Lichens are very sensitive to air pollution and with the advent of industrialization the lichens disappeared and the bare trees became blackened by sooth. Against this background the melanic form is less conspicuous than the light form. As a result the melanic variants were less easily caught by their predators than their lighter relatives and their share in the population increased rapidly.

Kettlewell has performed a series of experiments to support this explanation. In 1953 he released marked individuals of both variants in the polluted woods near Birmingham. About 90% of the natural population in this woods is melanic. More than twice as many dark as light variants were recaptured after one day in the wild. A repetition of this experiment in 1955 yielded approximately the same results. In the same year the experiments were also repeated in the unpolluted woods of Dorset, where the trees are covered with lichens and about 95% of the natural population is light-coloured. The results are the reverse of those of the Birmingham experiment: more than twice as many light-coloured as black-coloured forms were recaptured. In addition, in co-operation with Tinbergen, he observed the resting moths from a hide. In the Birmingham area they observed 58 moth being captured by birds, 43 of the captured moths were light-coloured and 15 melanic. In Dorset of 190 moth observed to be captured more than 86% were melanic and less than 14% light-coloured.

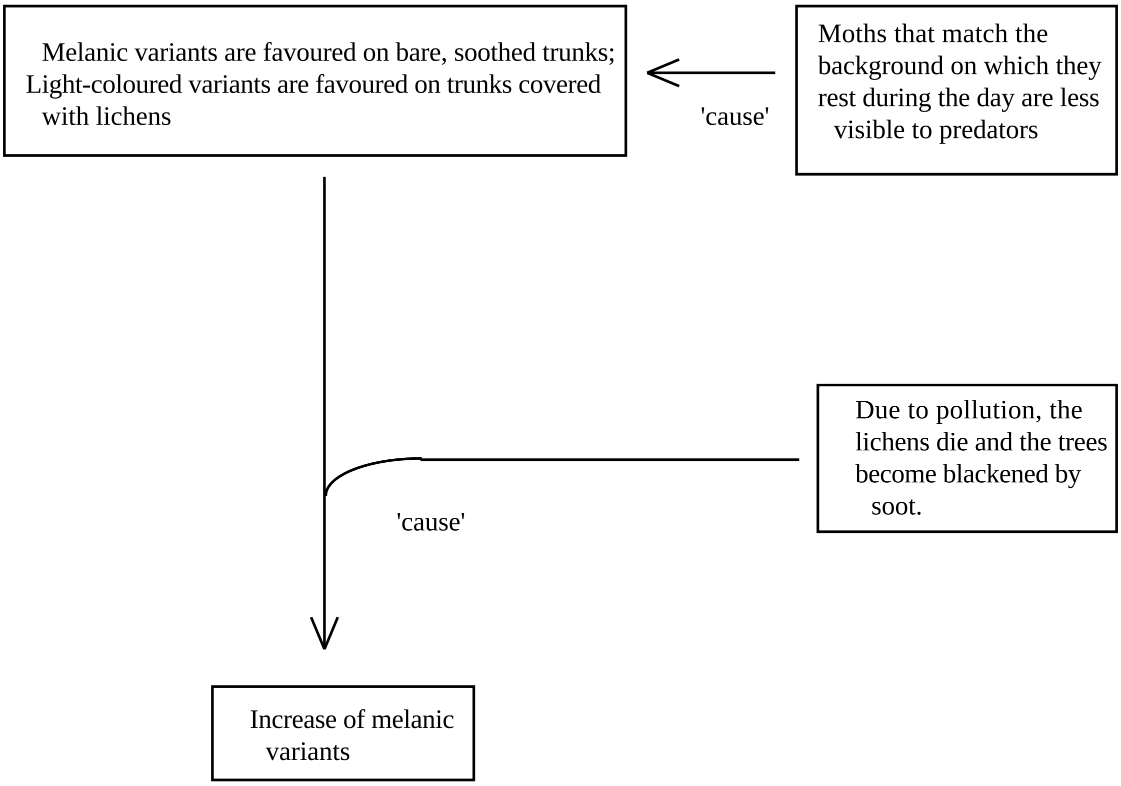

Fig 7.2: selection explanation of the evolution of industrial melanism

Hence, the phenomenon to be explained is the increase of melanic variants in the population in a certain period (see fig. 7.2). This phenomenon is explained by two factors, namely (1) differences in relative fitness between the melanic variants and the light-coloured variants in different environments (melanic variants are favoured on bare, soot-covered trunks and light-coloured variants are favoured on trunks covered with lichens), and (2) a change in the environment (due to pollution the colour of the trunk became bare and blackened by soot). The fitness differences in their turn are explained by the fact that moths with a pattern of colours that match the background on which they rest during the day are less visible to their predators. (This account of industrial melanism is derived from Kettlewell 1959, Kettlewell 1973, Keeton & Gould 1993: 480-481)

Kettlewell primarily explains a past increase of melanic variants in the population, but this secondarily explains why the wings of current moths have the colour pattern they have. That is it explains why certain items (the wings) have the character (colour pattern) they have.

Evolutionary selection explanations, as I have defined them, are concerned with changes at the population level. Do we need such an explanation to attribute a function according to the etiological theory? At least in Millikan's view the answer to this question is no. Millikan reminds us that selection might not only explain evolutionary change. It might also explain why a certain item or behaviour has not changed (has not evolved) in the recent past:

> If natural selection accounts for a trait, that is something that happened in the past, but that past might have been, as it were, "only yesterday". Indeed, *usually* the relevant past is "only yesterday": the *main business* of natural selection is steady maintenance of useful traits by continual selection against new intruders in the gene pool (Millikan 1989a: 173).

The idea is that in the course of time in a population there regularly turn up mutant variants in which the character of an item has changed in such way that that item performs a certain activity less well, for example, there regularly turn up mutants in which due to a change in the heart's structure the heart pumps the blood less well. These variants are subsequently weeded out by natural selection because they perform that activity less well, for example, because they pump blood less well. In those cases the character of an item or behaviour is maintained by selection and on Millikan's theory it is justified to call the activity concerned a proper function of that item:

> A trait's function is what it actually did — did most recently — that accounts for its current presence in the population, as over against *historical* alternative traits no longer present (Millikan 1989a: 174).[4]

---

[4] Note that in this paper Millikan attributes functions to traits. In her other work functions are attributed to items which is closer to the way biologists talk of functions.

It is useful to distinguish two kinds of selection explanations: evolutionary selection explanations and equilibrium selection explanations. Evolutionary selection explanations explain changes at the level of the population and equilibrium selection explanations explain the maintenance of a certain trait at a certain frequency in the population. Similarly it is useful to distinguish two variants of etiological theories: an origins variant and a maintenance variant. The origins variant defines the functions of a trait as the effects that caused the past increase of that trait in the population. The maintenance variant (endorsed by Millikan) emphasizes the effects that have maintained a certain trait at its present level in the population in addition to the effects that caused its original increase. On the origins variant to attribute a function to a trait one needs evidence that the trait increased in the population due to natural selection for performing that function, on the maintenance variant one needs evidence that in the recent past the trait has been maintained by natural selection for performing that function.

7.3 The proper place of proper functions

In this section I discuss the extent to which the etiological theory helps us to understand the use of function attributions in explanations in functional biology.

As I said Millikan on the one hand maintains that her definition is not meant to capture the biologist’s usage and on the other hand she gives the impression that she thinks that her definition applies to almost all function attributions in biology (among others because her examples of function attributions include standard function attributions in biology such as “a function of the heart is to pump the blood around”). Anyway Millikan does not give a detailed example of the way in which biologists come to attribute a function to an item or behaviour, neither does she study how those attributions are actually used in biology.

Neander’s theory is explicitly meant to capture the biologist’s usage. However, her main published defence of the etiological theory (Neander 1991a) deals with philosophical objections against the etiological theory. She does not give detailed examples and does not show in detail that her theory applies to biological practice. The lack of detailed examples makes it difficult to understand to what kinds of function attributions she thinks her theory applies and in what kind of explanations she thinks these function attributions are used. Anyway, as the above quotations show she thinks that her theory applies to at least two function attributions that on my definition are attributions of function as causal role (namely the attribution of the function to pump blood to the heart and the attribution of the function to assist in grasping objects to my opposable thumb).

I argue that the etiological theory is of no relevance to understand talk of functions in functional biology. I have three arguments:

1) The kind of evidence biologists bring up when they attribute functions to the parts, organs and behaviours of organisms is not the kind of evidence one would expect on the etiological account. This shows that the etiological theory does not apply to this kind of function.
2) The kind of functional explanations sketched by the proponents of the etiological account have only a remote resemblance to the explanations offered by functional biologists.
3) The interpretation of biological categories as categories that unite items and behaviours by their function is wrong.

The etiological theory is at best an account of the fourth notion of function (function as selected effect) and this notion of function does not play an important role in functional biology.

7.3.1 The evidence against proper functions

Although the meaning of a statement is not equivalent to its evidence, it seems that there must be a relation to what is asserted by means of a certain scientific statement and the evidence brought up to support that thesis: the relevant scientific community must think that the evidence gives them an indication of the truth of what is supported. But this means that the evidence brought up in support of a certain kind of thesis, gives us some indication about what is asserted. In other words, one need not be a verificationist to maintain that the evidence which is brought up in support of a certain kind of thesis provides an important way to evaluate whether a philosophical theory applies to theses of that kind. Neander, for example, maintains that the kind of conceptual analysis in which she is engaged

> involves trying to describe the criteria of application that the members of the linguistic community generally have (implicitly or explicitly) in mind when they use the term (Neander 1991a: 170).

If the etiological theory would apply to such function attributions as “the heart has the function to pump the blood around” and “the function of my opposable thumb is to assist in grasping objects” what kind of evidence would one expect for these function attributions? As selection consists in heritable fitness differences and results in , we might expect:

- (i) evidence that variation occurred in regard to the relevant trait,
- (ii) evidence of natural selection, that is evidence that
    - (a) in the environment in which those organisms lived organisms having the trait performed the function-to-be-attributed different from their competitors, and
    - (b) in that environment the way in which the function was performed by the organisms having resulted in better life chances than the way in which that function was performed in organisms lacking the trait,

(iii) evidence that this variation is heritable,
(iv) evidence that the predicted results occur.

Natural selection is not easy to detect. The authoritative guide to detecting selection in natural populations is John Endler's *Natural Selection in the Wild* (1986) . Endler discusses ten different kinds of methods to detect natural selection. Problems may result among others from statistical bias and from incomplete knowledge of genetics, population structure and ontogenetic development. Endler presents a lists of all known direct demonstrations of natural selection (p. 129-53). These demonstrations are distributed among 314 traits, and among 141 species. This is quite impressive as a demonstration of the occurrence of natural selection in the wild, but it fails as a justification of the common function attributions in biology. Moreover, most of these studies are concerned merely with demonstrating *that* natural selection occurred, but not with the *why* of selection. None of this studies is concerned with the function attributions philosophers usually mention as examples, such as the functions of heart and kidneys, the function of sperm and so on. This means that if the etiological theory applied to these function attributions biologists would have no support for their standard function attributions.

If one requires that a function attribution tells us something of the origin of the trait (rather than on the maintenance) the problems increase. Amundson & Lauder (1994) have spelled out these difficulties. I quote from this paper:

> There are enormous practical difficulties in determining just what the selected effect of a structure was in the first place. Many structures are ancient, having arisen hundreds of millions of years ago. During this time, environments and selection pressures have changed enormously. How are we to reconstruct the ancient selected effect?

If there is enough palaeontological information, it is sometimes possible with help of physical theories to estimate the likely action of selection in the past and to construct a possible scenario of the evolution of a certain trait, but many structures, in particular in fossils, are not amenable to such an analysis.

More specifically, on Millikan's account (which emphasizes selection for maintenance of a trait in the recent past) one would expect that someone who wants to attribute the function to pump the blood around to the heart comes up with (i) a list of deviations that regularly occur in the heart, (ii) an account which shows that (a) these deviations result in differences in the way in which the blood is pumped, and (b) the life chances of variants which deviate hearts are lower than that of normal organisms due to the difference in the way in which the blood is pumped, (iii) evidence that the deviations are heritable, (iv) an analysis which shows that the percentage of people having the deviant heart decreases with the age (or other evidence of the predicted outcome). I could not find anyone providing this kind of evidence for a function attribution. To avoid misunderstandings, my point is not that this kind of evidence does not exists.[5] Rather my point is that nobody cares to present this evidence in favour of the bulk of function attributions, even if it exists. This is a strong indication that the etiological theory does not apply to those function attributions. When biologists identify functions they seldom delve into the selection history of an item, behaviour or trait. Instead they try to figure out (i) what role a certain part, organ or behaviour has in the maintenance of an organism (in which case the study results in an attribution of a causal role), or (ii) what the survival value is of a certain trait as compared to another trait (in which case the study results in a claim about survival value). I show this by means of some examples.

My first example concerns the attribution to the heart of the function to pump blood. The evidence for this attribution was provided by Harvey in the beginning of the seventeenth century (example 2.1 of section 2.2.2). Harvey's main point in regard to the heart was that the heart movement does not originate from the chest and the lungs but from the heart itself and that the active phase is the expansion phase rather than the contraction phase. Harvey discovered that the heart of some amphibians and reptiles remain active after killing the organisms and used observations of how the heart works in those organisms to argue for his thesis that the function of the heart is to pump blood. He also observed that the heart is hollow muscle. Of course, there is no hint of selection in Harvey's argument.

Neander (1991a) dismisses appeal to Harvey on the ground that her etiological theory is meant as an analysis of present day biology and that the notion of function has changed with the arrival of Darwin's theory. I think she is wrong. The present day attributions of the function to pump the blood to the heart derive from Harvey. Current textbooks on morphology hardly give evidence for function attributions. Evidence is presented only in the rare cases in which the attribution is controversial. In the case of the heart, textbooks confine themselves to saying that the function of the heart is to pump blood without presenting the evidence for that attribution (the function attribution is used to explain such things as how the heart and the circulatory system work, why the heart has the structure it has, how it developed in the course of the evolution (see section 2.3.3)). This means that the attribution is entrenched in morphology to such an extent that textbooks state the function without giving evidence. The attribution became

---

[5] One should not underestimate the difficulty of providing such evidence. For example, in the case of the heart it is not enough to point to the fact that in many hospitals all over the world there are people with heart diseases and that these diseases involve the capability to pump blood. One must show that these diseases have a genetic component, that they influence fitness (it might for example be the case that most heart diseases occur at an age at which they do not influence the expected number of offspring) and it would be nice to have also some evidence that the predicted result of nautral selection occurs. The latter kind of evidence is, perhaps, the most difficult to obtain (as far as I know there are no means to detect deviant hearts at an early age).

entrenched in biology due to Harvey's work. If Neander is right that the meaning of the term function changed with Darwin one would expect that after the arrival of Darwin's theory biologists would point out that the concept of function has changed and that for that reason Harvey's evidence was bad evidence. No such thing happened. Indeed, if any evidence is presented today for attributing the function to pump blood to the heart it is Harvey's evidence. For example, I first learned the evidence for this function attribution, from a movie about Harvey shown as part of a practical course in zoology.

However, I need not appeal to Harvey to argue that the evidence presented in favour of the common function attributions in morphology is not of the kind one would expect if the etiological theory would apply to those attributions. In 1961 Jacques Miller showed that the thymus had a function in the development of the immune system (example 2.2 of section 2.2.2). His evidence consists of the results of extirpation experiments in new-born mice. There is no investigation into the history of the thymus, no list of common mutants, no argument that the mutants are worse off than the normal forms, no investigation into heredity and no investigation of natural populations which could affirm a selection hypothesis. In sum: the evidence Miller provided in support of the hypothesis that the thymus has a function in the development of the immune system is evidence that the thymus is (in current organisms) involved in the development of the immune system not evidence that the thymus was selected for that role.

Especially instructive is my third example: Kurt Schwenk's (1994) argument for the attribution of trail-following function to the tongue of snakes (and some other reptiles) (example 2.3 of section 2.2.2). As I said in section 2.2.2 his evidence consists of: (1) observations of the behaviour of snakes during trail following, (2) experiments showing that removal of the forked portion of the tongue eliminates the ability to follow scent trails, but not the delivery of stimuli to the vomeronasal organ, (3) comparative data showing a rough correlation between depth of tongue bifurcation and ability to follow chemical trails, (4) observations of the nervous system showing circuits suitable for comparing signal strength from each side of the tongue. There is no appeal to history nor to selection in this part of the argument. The example is especially instructive because Schwenk discusses the selection history of the trait in addition to its function (albeit superficially). At this point he does present historical evidence: he fits the trait 'having a forked tongue' on a phylogenetic tree and shows that the forking is a derived character which emerged at least twice in the evolutionary history. He then speculates that the forking might have arisen because it conferred a performance advantage in chemical delivery to the vomeronasal organ and that subsequently the forking increased due to selection to perform this role better.[6] Whereas Schwenk maintains that his conclusion about the function of the tongue

---

[6] This is merely a speculation because Schwenk does not tell us what the performance advantage was and he does not attempt to correlate the emergence of a forked tongue with the need for such a performance advantage.

in trail following is “inevitable” (p. 1574), he is pretty uncertain about why the tongue became forked (at p. 1567 he maintains that “the reason for its initial evolution remains obscure”. “Probably [...] it did initially confer a performance advantage in chemical delivery to the vomeronasal organ.” Subsequent selection for increased efficiency in scent trail-following “may have caused” the tongue to become increasingly forked). The facts that (i) the evidence which Schwenk presents in support of his function attribution differs from the evidence he presents for supporting his hypotheses about selected effect, and (ii) Schwenk is confident about the function attribution and hesitant about the selected effects, support my view that the etiological theory does not apply to the common function attributions in morphology.

My first three examples concerned examples in which the evidence provided in support of a function attribution shows that the attribution was concerned with the causal role of the item in study not with the effects for which that item was selected. Let us now look at an other kind of examples

My fourth example concerns the foraging behaviour of black headed gulls (example 3.1 of section 2.2.3). As I said, Gotmark, Winkler and Anderson (1986) argue that this behaviour has the function to enable each gull to catch more fish than it would have done if it fished alone. They support this attribution by means of a laboratory experiment which shows that gulls in groups catch more fish (per gull) than do gulls that fish solitary and that the number of fish catched per gull increases if the group size increase. I admit that this kind of evidence comes closer to the kind of evidence one would expect if the etiological theory applied (closer than the evidence in my first three examples). The experiments support the view that the fitness of a gull that tends to flock is greater than the fitness of a gull that tends to fish alone. However, these experiments do not provide sufficient evidence for the view that there recently was selection. They do not show that there regularly turn up individuals that tend to fish alone, which survive less well due to this tendency and they do not show that this tendency is heritable.

Similar considerations apply to my fifth example: the egg shell removal behaviour of black headed gulls. The experiments of Tinbergen and his colleagues (1962) show that the actual birds that remove the egg shell would be favoured by selection over hypothetical variants that would not perform this behaviour. According to Millikan's etiological theory, however, functions are attributed not on the basis of “what would happen if”, but on the basis of what actually happened in the past. This requires evidence that the hypothetical selection established by the authors has actually played a role in the evolution or maintenance of the egg shell removal in the past. It should for instance be shown that every now and then a mutant gull turns up that has decreased fitness because it does not remove the empty egg shell, or waits to long before doing so, or dumps the empty shell at a too short distance. The researchers do not provide evidence of this kind. Yet, they take it that their studies establish the function of the trait concerned.[7]

I have presented five examples in which the evidence provided by the authors differs from the evidence one would expect on Millikan’s theory. This may either mean that the authors are over-confident or that they use the word 'function' in a sense different from that of Millikan. In my examples we can safely exclude the first possibility. The examples are either derived from renowned refereed journals such as *Science* and *Nature* (Schwenk, Gotmark, Winkler & Anderson) or else they have become classics in their field (Harvey, Miller, Tinbergen). The conclusion that they use the term 'function' in a sense that is different from Millikan's seems justified. In the first three examples the term 'function' refers to the causal role of a certain part or organ, in the last two examples the term 'function' refers to the survival value of a behavioural pattern. I do not deny that the word 'function' is sometimes used by evolutionary biologists in Millikan's sense of proper function. My point is, rather, that functional biologists often use the term ‘function’ in a different way and that no account of function in biology is complete if it ignores that other use.

7.3.2 Functional explanations

In the preceding section I argued that the kind of evidence put forward in support of the standard function attributions in functional biology is not of the kind one would expect on the etiological theory. This shows that the etiological theory does not apply to these kind of function attributions. In this section I am concerned with the question whether or not the etiological theory offers insight in explanations in functional biology.

In chapter 2 I argued that functional biologists aim to explain the way in which an organism is built and the way it behaves. They do so in four complementary ways: (1) physiological explanations, (2) design explanations, (3) developmental explanations, (4) evolutionary explanations. I also argued that functions as causal roles have a key role in three kinds of

---

[7] The point that evidence of survival value is not evidence of selection has been made emphatically by Tinbergen (1963) and by Hinde (1975). Tinbergen emphasizes the difference in method between establishing survival value and establishing selection. Survival value is, according to him, determined by means of experiments which allow us to attribute survival value beyond any reasonable doubt. Evidence of selection, on the other hand, is provided by means of comparision and hypotheses about selection are rather speculative. Hinde emphasizes that comparision with a hypothetical organism does not establish selection. Hinde distinguishes between strong and weak meanings of function. Function in a strong sense corresponds to my function as selected effect, function in a weak sense to my function as survival value. Functions in the weak sense are established by comparing the real organism with a hypothetical one, functions in the strong sense are established by comparing real competitors. The latter kind of evidence is, according to Hinde, much more difficult to supply.

explanations: capacity explanations (which are a special kind of physiological explanations), design explanations and evolutionary explanations. In this section I am concerned with the question what insights the etiological account provides in these kinds of explanation and whether my account should be revised in the light of the etiological theory.

As I said in the introduction to this chapter, the etiological theory is primarily a theory of function, not of functional explanation. Millikan and Neander, the main proponents of the etiological theory, pay little or no attention to explanation. Millikan makes some scant remarks on "functional explanations" in her "An Ambiguity in the Notion 'Function'" (1989a) and Neander devotes a paper to the alleged feature that attributions of proper functions generate what she calls "teleological explanations" (Neander 1991b). Neither Millikan nor Neander discuss a detailed example of an explanation in biology.

Although it is clear that on the etiological account function attributions are singled out by their role in selection explanations, it remains unclear what according to the proponents of the etiological theory the relation is between on the one hand selection explanations and the resulting attributions of "proper" functions and on the other hand the explanations which biologists call 'functional explanations' (and which I call 'design explanations'). Would the proponents of the etiological account agree with me that design explanations are a different kind of explanations beside physiological explanations (among which capacity explanations), developmental explanations and evolutionary explanations (among which evolutionary selection explanations)? If so, would they agree with me that these explanations appeal to causal roles and survival values or would they maintain that these explanations appeal to proper functions in their sense? Would they maintain that design explanations appeal to past selection in some way or other? What is the structure of such explanations on their account?

Millikan

In her 'An Ambiguity in the Notion "Function"' (1989a) Millikan scorns Horan (1989) and Bigelow & Pargetter (1987) for confusing "two types or senses of 'function' and 'functional explanations'" (p. 172). The two kinds of function are on the one hand "Cummins-style functions (Cummins 1975) " and on the other hand "selected functions" (or "proper functions" — Millikan 1984, 1986, 1989b). "Cummins-style functions" correspond to the kind of function I call "function as causal role" (function$_2$); "selected functions" correspond to the kind of function I call "function as selected effect" (function$_4$). According to Millikan both kinds of function can be used to explain the presence of the item to which the function is attributed, but the resulting kinds of explanations should not be confused. Millikan distinguishes two types of functional explanations: "functional explanations in Cummins sense" (p. 175) and "functional explanations that make reference to natural selection" (p. 174). "Functional explanations in Cummins sense" correspond to the kind of explanation I call "capacity explanation", "functional explanations that make reference to natural selection" correspond to the kind of explanation that I call "selection explanation". Millikan maintains that functional explanations that make reference to natural selection appeal to functions as selected effect (and not to causal roles), whereas causal roles are used in functional explanations in Cummins's sense (that is in capacity explanations).

Functional explanations that appeal to natural selection explain the presence of a certain trait simply by saying that that trait has a certain selected function, say *F*. Millikan emphasizes that on her theory to say that a certain trait has a function *F* is just another way of saying that in the past that trait was selected over its competitors because it did *F*.

> To say that a trait *T* has the proper function *F* is to say that *T* had a history during which it was selected for doing *F*. So if you want to know why current species members have *T* the answer is, very simply, *because T has the function F,* that is, because *T was* selected for because it *did F*. *F* was not of course the function of the very first tokens selected, even though they did *F* and that is why they were selected. They were not selected "because they had functions" but because they did *F* (Millikan 1989a: 174).

As I discussed in chapter 5, Cummins (1975) maintains that attributions of causal roles have no place in selection explanations. Millikan agrees with him:

> It is important to see that Cummins' sense of function does not fit with the theory of natural selection to help explain the current presence of a trait in a species (Millikan 1989a: 175).

Only attributions of functions of selected effects may be used in selection explanations:

> [selected function] is the only kind of function that is well defined and that can enter into functional explanations that make reference to natural selection (Millikan 1989a: 174).

Although Millikan agrees with Cummins that causal roles may not be used in selection explanations, she argues, *pace* Cummins, that there are cases in which the presence of an item might be explained by means of a functional explanation in Cummins sense (and hence by appeal to the causal role (Cummins function) of the item to which the function is attributed). Such cases concern causal roles which help to explain the reproduction of cyclical system (such as "the life cycle of a species" (p. 176)). In such case the presence of a certain item in the current phase of the cycle might be explained by the fact that that item performed its causal role in the preceding phase. For example, the presence of my heart might be explained by the fact that the hearts of my parents performed their causal role (in "the life cycle of the species") to propagate the blood.

> So the Cummins functions of various features of a biological species, when these functions are defined relative to the life cycle of the species, do help to explain the most recent genesis of these various features as present in current individuals (Millikan 1989a: 176).

Millikan's concludes by saying that as (i) most selected functions coincide with a Cummins function that helps to explain "how a species [sic!] has recently been surviving, developing and reproducing" (p. 175)[8] and (ii) both kinds of functional explanations may explain the presence of the item to which the function is attributed the two sense of function and functional explanation are easily confused.

As I have shown in chapter 6, Horan and Bigelow & Pargetter are interested in a third type of function, namely function as survival value (function$_3$) which they confuse with function as causal role and erroneously take for the one and only kind of function. Millikan too falls prey to confusion, not only because she fails to see that Horan and Bigelow & Pargetter are interested in a third kind of function, but also because she has a mistaken idea of the explanatory use of functions as causal roles.

Biologists do not use causal role functions to explain the presence of the item to which the function is attributed in the way described by Millikan. It is not clear what Millikan means by "the life cycle of the species" but, of course, someone may explain the presence of my heart by appeal to the fact that the hearts of my parents performed their causal role in the development and reproduction *of my parents* but this has nothing to do with the way in which biologists use the notion of causal role in explanations. In this kind of explanation, the fact that the explanation appeals to the causal role of the item the presence of which is to be explained is accidental. One may just as well explain the presence of my heart by appeal to the fact that my parents *lungs* performed their causal role. This in contrast to the way in which biologists appeal to causal roles in explanations. As I have shown by means of many examples in several chapters, biologists use attributions of causal roles to explain (i) how a certain organism, item or behaviour is able to perform a certain task, (ii) why it is useful to certain organisms that they have a certain item or behaviour, (iii) why it is useful to certain organisms that a certain item or behaviour has the character it has, and (iv) how that item or behaviour got its character in the course of the evolutionary history. In a certain sense of 'why' the explanations ad (ii)-(iv) may be taken as explanations of why a certain trait is present in a certain type of organisms. Explanations ad (ii) (which answer a type 4a question) explain why$_2$ certain organisms have a certain item or behaviour. They do so by telling us why the causal roles performed by that item or behaviour are useful to those organisms. Explanations ad (iii) (which answer a type 4b question) explain why$_2$ a certain item or behaviour has the character it has (if one thinks of the character

---

[8] To my mind the very notion of a Cummins function "in the life cycle of species" is confused. It is individuals not species that survive, develop and reproduce. Perhaps Millikan means to say "how the member of a species have recently been surviving, developing and reproducing". Cummins defines functions as capacities of subparts that help to explain the capacities of the part to which the subpart belong. However, the things certain individuals did that contributed to their survival, development and reproduction are activities not capacities.

of an item as a trait one may phrase this as why$_2$ that trait is present). They do so by telling us why the causal role of that item or behaviour is better performed if the item or behaviour in question has the character it has than if it had some other character. Explanations ad. (iv) (which answer a type 6 question) tell us why$_3$ a certain item or behaviour has the character it has. They do so by telling us that the item or behaviour in question acquired its character due to the fact that in an ancestral population variants in which the item or behaviour in question had the character it now has were fitter than variants in which that item or behaviour had another character because in the circumstances in which those organisms lived the causal role was better performed by an item or behaviour with the present character than by an item or behaviour with that other character. Note that in all these kinds of explanations the appeal to causal roles of the item in question is essential. One does not explain why the heart is useful by appeal to the causal role of something other than the heart (such as a lung), nor does one explain why it is useful that the heart has the structure it has and why it acquired that structure in the course of the evolution by appeal to the causal roles of other organs.

Millikan’s idea that causal roles may not be used in selection explanations is equally mistaken. Consider, for example, Abramson's (1993) study of the structure and evolution of dentition of true lemmings. Lemmings are small rodents that inhabit the tundra's of Scandinavia, Siberia and North-America. There are two subgroups: true lemmings (*Lemmini*) and collared lemmings (*Dicrostonyxini*), both belong to the subfamily of vole (*Arvicolinae*). The structure, trends in variation and rate of evolution of dentition in true lemmings is remarkably different from that of collared lemmings and, indeed, of all other voles. Abramson argues that these differences “can be understood from the quite different functioning of the masticatory apparatus” (p. 687). True lemmings differ in their food preferences from all other vole. They mainly eat graminoids and mosses, which are poor in mineral nutrients and contain a large volume of hardly digestible hemicellulose fibres. This diet imposes demands on the masticatory apparatus which are different from those of other voles with another diet:

> Thus the masticatory apparatus of the Lemmini *must* deal with the consumption of far greater quantities of food per unit of time than is the case with voles, containing in addition much higher amounts of hardly digestible cellulose (Abramson 1993: 696, emphasis mine).

These different demands explain both the differences in structure and the differences in evolution. In true lemmings the greater part of the cutting edges of the upper and lower molars form reciprocally curved blades. As a consequence some blades meet along the direction of movement. In collared lemmings, the cutting edges of the upper and lower molar meet obliquely. The arrangement in true lemmings increases the contact area (in comparison to that of other voles). Such curved blades entrap food better, which allows the processing of a larger amount of food per unit of time. In the conclusions section this point is expressed by means of a functional counterfactual:

> It has been shown that with the arrangement of cutting edges characteristic of highly evolved voles, food with a high concentration of cellulose will slip off the masticatory surface instead of being cut (Abramson 1993: 697)

The arrangement of cutting edges in its turn explains the arrangement of the muscles: the arrangement of cutting edges in true lemmings requires an increase in the force normally applied to the occlusal surface and the muscles differ in accordance with this requirement.

The evolution of tooth patterns in vole is a well studied area. The main trends observed in many different lineages are: acquisition of additional triangles (dentoenamal elements) on the rear end of the lower molars and the front end of the upper ones, together with an increase in the degree of alteration of salient angles. These trends are extremely clear in the evolution of collared lemmings. True lemmings, however, show a different trend, namely an increase in the relative width of the molars and a tendency to fuse the triangles. Roughly spoken the trend is towards flattened molars. Both in true and in collared lemmings these trends result from selection for achieving maximum efficiency of the molars. However, due to the dietary differences the means to increase that efficiency differ in these two lineages.

> The trend leading to one solution or another depends greatly on the genetic base, and, to a very significant degree, on demands imposed on the masticatory apparatus by the mechanics connected with feeding specialization (Abramson 1993: 695/6).

The flattening of the molars is a means to achieve greater efficiency in herbivores with a low caloric diet.

> In Lemmini, with their characteristic chewing movements and the nature of occlusion, increased mastication efficiency is achieved not with the elongation of the cutting edges by the acquisition of new elements, but by the increase in the relative width of the teeth and the fusion of the alternating triangles (Abramson 1993: 693).

In sum:

> The distinctive means of increasing the efficiency of the masticatory apparatus seen in Lemmini not only reflects definite genetic history but can also be explained by their very unusual food specialization and the demands it imposes on the mechanics of their masticatory apparatus (Abramson 1993: 687)

(The remarks about genetics appear out of the blue and remain unclear).

Note, that Abramson's explanation of the structure and evolution of the molars of true lemmings proceeds in the same way as Schwenk's (1994) explanation of the forked character of the snake's tongue. Both explanations start with the attribution of a causal role to the items in study. This attribution is subsequently used in both a design explanation and an evolutionary explanation of the structure of the item in study. The design explanation shows that the relevant causal role is better performed if the item has the structure it actually has than if it has another conceivable structure. The evolutionary explanation points out that a certain trend in evolution can be explained as the result of selection of variants that performed the relevant causal role more efficiently in the conditions that apply to the organisms involved. In Schwenk's case the attribution of a causal role in chemosensory tropotaxis to the snake's tongue takes up the large part of the paper; in Abramson's case the attribution to the molars of a causal role in mastication is well established and the explanatory part gets the most attention. Schwenk's design explanations explains why it is useful that the snake's tongue is forked (why$_2$ the tongue is forked) by pointing out that blunt tongues are of no use in chemosensory tropotaxis. Abramson's design explanation points out that given the dietary preferences of true lemmings the causal role of the molars (mastication) is better performed if the molars have the structure they have in true lemmings than if they had the structure they have in collared lemmings. Schwenk's evolutionary explanation suggests that the tongue of snakes became increasingly forked in the course of evolution due to selection of variants in which the tongue performed its causal role in chemosensory tropotaxis more efficiently. Abramson's evolutionary explanation explains the evolutionary tendency toward flattened molars in the lineage of true lemmings as the result of selection of variants in which the molars performed their role more efficiently. This explanation starts with the observations that the molars have a causal role in mastication and that the diet of true lemmings consists of graminoids and mosses. It then points out that if the diet consists of graminoids and mosses the molars perform their causal role more efficiently if they are more flattened. For that reason in the course of evolution variants with more flattened molars were favoured again and again over variants with less flattened molars, which resulted in the tendency to flatten the molars.

The appeal to causal roles in selection explanations such as the ones provided by Schwenk and Abramson shows that Millikan's idea that selected functions are the only kind of functions that can enter into selection explanations is mistaken. Indeed, when biologists talk of function in the context of selection explanations they seldom use that term in a historical sense—if ever. I have already quoted Millikan explaining how attributions of selection functions are used to explain the presence of a certain trait:

> if you want to know why current species members have T the answer is, very simply, because T has the function F, that is, because T was selected for because it did F (Millikan 1989a: 174).

The way in which biologists talk is different in a subtle, but fundamental way. Biologists often say that a certain structure (e.g. flattened molars) was favoured by selection because it increased the efficiency with which a certain function (e.g. mastication) was performed. For example, according to Abramson the flattened molars were selected because they increased the efficiency of the masticatory apparatus. This means that the selected function (in Millikan's sense) of the flattened molars is to increase the efficiency of the masticatory apparatus. But no biologists will say that ‘‘increasing the efficiency of the masticatory apparatus” is a function of the flattened molars. They say that flattened molars were favoured by selection because they increased the efficiency with which the mastication function is performed, but they do not say that “increasing the efficiency of the masticatory apparatus” is a function. The function to which they appeal is mastication (which is a causal role), not increasing the efficiency of a certain item; and this function is attributed to the item (the molar) the structure (flattened) of which is to be explained not to that structure itself. This means that their use of the term ‘function’ does not conform to the etiological theory.

Moreover, these examples show that selection explanations add more to our knowledge than Millikan seems to think. Abramson does not explain the flattened molars *simply* by saying that they were selected because they increased the efficiency of the masticatory apparatus. One of the main points of his explanation is that more flattened molars are more efficient than less flattened ones because of the demands imposed upon their function (mastication) by the kind of food consumed by the true lemmings. Millikan’s idea of selection explanations as consisting of simply an attribution of a selected function completely negates this part of selection explanations.

Let me draw some conclusions. Millikan (1989a) distinguishes two kinds of functional explanations that might explain why certain organisms have a certain item: functional explanations in Cummins sense and functional explanations that make reference to natural selection. I have argued (i) that the way in which biologists use attributions of causal roles in design explanations to explain the presence of a certain trait has nothing to do with Millikan’s “functional explanations in Cummins sense”, and (ii) that Millikan’s account of functional explanations that make reference to natural selection leaves out of sight much of what is achieved by such an explanation because it fails to take into account the attributions of causal roles that are an essential part of such explanations. The conclusion is that Millikan’s (1989a) remarks on functional explanations are of no use if one wants to understand explanation in functional biology.

Neander

Millikan is not the only one who seems to think of functional explanations as consisting of a single attribution of a selected function. This idea seems akin to the etiological approach. Larry Wright, for example, maintained that the “insight” that functional explanations consist of a single function attribution as the foundation of his theory of function, which became a predecessor of several of the current etiological theories, among which is that of Neander. According to Wright the key to a satisfactory account of function is the insight that

> merely saying of something, X, that it has a certain function, is to offer an important kind of explanation of X (Wright 1973: 154)

an insight that all other theories of function have failed to take into account.

Neander too gives the impression that she thinks of functional explanations as one sentence attributions of selected functions. In her "The Teleological Notion of Function" (1991b) she maintains that function attributions "universally and intrinsically" "generate" or "justify" "teleological explanations". As an example of such an explanation she mentions an explanation of why penguins are myopic on land which points out that this is a "by-product of an optical system that has the primary function of providing sharp visual focus under water where the penguins find their food" (p. 454). This explanation is, in her view, teleological because it appeals to what the eyes are for:

> The function of the penguins' eyes of providing sharp visual focus under water explains why the penguins have the eyes they have, and why they have land myopia, by explaining what their particular optical system *is for* (Neander 1991b: 454. emphasis hers).

The terms "generate" and "justify" suggest that, according to Neander, function attributions and functional explanations are different things. But the differences as well as the relations between the two remain unclear. How is a teleological explanation generated from a function attribution? What is the structure of the resulting explanation? Neander contrasts her view that "function attributions universally and intrinsically justify teleological explanation" (p. 458) with (i) the views of Beckner (1959) and Cummins (1975) that "any so-called 'teleological explanation' derived from biological function is merely misguided pseudo-explanation" (p. 458), and (ii) the views of Canfield (1966), Ruse (1973) and Wimsatt (1972) who, according to Neander, maintain that "function attributions usually (but not always) justify teleological explanations, and only when combined with background theory extraneous to (not implied by or contained in) the function attribution" (p. 458).[9] Neander's main argument against these views rests on her judgement that the idea that function attributions are explanatory is "quite robust" (p. 457). According to her, in the past function attributions were standardly interpreted as stating the reasons why God gave a certain organism the item to which the function is attributed.[10] Notwithstanding the insight that "in modern biology" "we cannot properly inter-

---

[9] Neander's characterization of the position of Canfield, Ruse and Wimsatt is slightly inaccurate. They do not say that function attributions "usually (but not always) justify" such explanations. Rather they say that function attributions are sometimes used in functional explanations (but function attributions have other uses too) and that if function attributions are used in functional explanations they are combined with background knowledge.

[10] Neander does not substantiate this claim, with which I disagree. Perhaps, Neander describes the way in which natural theologists in the first half of the nineteenth century viewed function attributions but natural theology is

pret” function attributions in this way, function attributions are used again and again to explain why certain organisms have the item to which the function is attributed.

> That the koala’s pouch has the function of protecting its young does seem to explain why koalas have pouches. That the bee’s dance is for directing other bees to pollen does seem to explain why bees dance (Neander 1991b: 457)

According to Neander it is “just barely possible” that this apparent explanatory power is illusory and an account which shows that such explanations are scientifically legitimate is preferable to an account that maintains that “we” are persistently irrational. The etiological theory provides such an account and is for that reason to be preferred over the views she opposes.

Neander is right that philosopher’s should not explain away established scientific practice as irrational. However, her representation of that scientific practice is mistaken. Neander does not substantiate her claim on penguin myopia, koala pouches and bee dances with reference to scientific literature. This makes it difficult to evaluate her representation of scientific practice in these cases. However, in section 6.6.2 I showed by means of several other examples that the explanations which biologists call ‘functional explanations’ (and which I have called design explanations) do not explain why certain organisms have a certain item or why an organism performs a certain activity simply by citing the function of that item or activity.[11] Design explanations come in two kinds. One kind, answers type 4b questions: they explain why a certain items or behaviour has a certain character. The other kind answers type 4a questions: they explain why an organism has an item that performs a certain task. Both kind of design explanations start by attributing a causal role (function$_2$) to an item. But this attribution is the first step in the explanation, not the whole explanation. To yield a design explanation the function attribution is explicitly combined with other statements. This shows that, contrary to what Neander claims, an account according to which function attributions are explanatory only when combined with other statements represent scientific practice better than an account according to which function attributions are intrinsically explanatory.

Perhaps, the idea that biologists routinely explain the presence of an item or behaviour by uttering a sentence of the form ‘the function of ... is ...’ rest on a failure to distinguish attribu-

---

only a minor influence in the history of biology. In the traditions of Aristotle, Cuvier and Von Baer functions were interpreted as roles in the maintenance of the organism.

[11] In regard to Neander’s examples I’m pretty sure that the explanations offered by biologists are much more complex than Neander maintains. For example, to explain the penguin’s myopia it is not sufficient to point out that the eyes have the function to provide sharp visual focus under water. In addition, one should connect the myopia to that function for instance by explaining why an eye that provides sharp visual focus under water does not provide sharp visual focus on land.

tions of causal roles (function$_2$) from attributions of survival value (function$_3$). Neander's examples are examples of attributions of causal roles. Attributions of causal roles might be converted to the standard form 'the function$_2$ of item *i* is to do *f*' or 'item *i* has the function$_2$ to do *f*' but they are not explanatory without any addition. To yield design explanations function$_2$ attributions are either combined with statements which say why it is useful to the organism to perform role *f* or with statements which say why it is useful that an item that has role *f* has the character which item *i* has. Attributions of survival value are design explanations but they have a much more complex structure than attributions of causal roles and they can not be converted without loss of information to a sentence of the form 'the function of .... is ...'.

Kuipers & Wisniewski

A detailed account of the structure of functional explanations in the spirit of the etiological theory comes from Kuipers and Wisniewski (1994). In a series of papers Kuipers (1985, 1986b, 1986a, 1996) has distinguished two main kinds of explanation: 'explanation by subsumption under a law' and 'explanation by specification'. The latter fall into three types: functional explanations, intentional explanations and explanations by specification of a cause. Kuipers and Wisniewski (1994) outline the common structure of explanations by specification (discovered by Kuipers) in terms of the logic of questions developed by Wisniewski. For obvious reasons, I will restrict myself to their account of functional explanations.

Kuipers and Wisniewski's main example of a functional explanation in biology is "the biological function of the systematic fanning movement of stickleback is to supply the eggs with oxygen". Their analysis consists of two parts: an analysis of the meaning of function attributions and a reconstruction of the train of thought underlying the search for functional explanations. Kuipers & Wisniewski's analysis of the meaning of function attributions (they call them "specific functional statements") falls within the scope of an etiological theory. According to this analysis function attributions of the form "the biological function of trait β of organisms of type α is γ" state that:

(a_1) β of α is a positive causal factor for γ,

(a_2) γ is a positive causal factor for the reproduction and survival of α,

(a_3) both β and γ were causally, i.e. evolutionary effective, for α having β (Kuipers & Wisniewski 1994: 382).

The train of thought underlying the search for a functional explanation starts with an explanation-seeking question of the form 'why do organisms of type α have trait β?" (e.g. "why do male sticklebacks show systematic fanning movement?"). As a first step in answering this question one accepts as a hypothesis to be tested an "unspecified functional statement" of the form "trait β of organisms of type α is functional" (e.g., the systematic fanning movement is functional”). The main point of the functional explanation is to establish a specification of this unspecified statement in the form of a function attribution of the form “the biological function of trait β of organisms of type α is γ” (e.g. “the biological function of the systematic fanning movement of sticklebacks is to supply the eggs with oxygen”). Such a function attribution provides a direct answer to the question “what is the biological function of trait β of organisms of type α?”.

According to Kuipers & Wisniewski this function statement might also be seen as an answer to the original explanation-seeking why-question:

> Each direct answer to a question of the form [what is the biological function of trait b of organism of type a?]12 may be regarded either as an answer to the corresponding question of the form [why do organisms of type a have trait b?] or as a sentence which entails such a statement (Kuipers & Wisniewski 1994: 384).

Kuipers & Wisniewski do not provide an argument for this thesis, but it is clear that they think that the specification of a function in itself accounts for the presence of the trait to which the function is attributed.

Kuipers & Wisniewski do not give references to biological literature. However, one might expect that their example of a functional explanation (“the biological function of the systematic fanning movement of stickleback is to supply the eggs with oxygen”) refers to the experiments Kristensen performed in the 1940s. Let us have a look at this example.[13] Male sticklebacks build a tubular nest and guard it with a complex pattern of behaviour. They alternate periods of swimming around the nest with periods as long as 30 seconds in which they stay before the nest in a slanting position, head down, moving their fins in a quick regular rhythm. Kristensen performed a series of experiments which showed that this behaviour has the causal role to supply the nest with oxygen. He showed that the eggs die if the male is removed from the nest and also if the nest is shielded from the fanning male with a watch glass. However, if oxygen rich water is directed to the nest by means of a tube, the eggs survive the removal of the male. Ventilation is needed because of the tubular nest, fish species who lay their eggs on leaves in running water do not need to ventilate the eggs. This example shows that Kuipers & Wisniewski are mistaken at two points. First, Kuipers & Wisniewski’s analysis of the meaning of biological function errs in defining biological function in evolutionary terms. Kristensen is concerned with the role and survival value of the fanning behaviour not with its evolution. Second, Kuipers & Wisniewski account of only a part of the insight provided by Kristensen’s explanation. They ignore the part which explains the need for ventilation by appeal to the tubular form of the nest. The question ‘why do male sticklebacks show systematic fanning move-

---

[12] I have substituted the formulae in Kuipers & Wisniewski's quote by appropriate sentences

[13] As I couldn’t find the original literature I use Tinbergen’s (1976: 12) account of Kristensen’s experiments.

ment?' is a typical type (4a) question which is answered by a design explanation. The attribution of the function to supply the nest with oxygen is but the first step in answering the question. In addition Kristensen explains why this function is needed by appeal to the tubular form of the nest.

Conclusion

The accounts of functional explanation offered by Millikan, Neander and Kuipers & Wisniewski are inspired by a picture of functional explanations as consisting of a single function attribution in answer to a question of the form 'why do .... organisms have [trait] ....?'. However, the explanations put forward by functional biologists seldom consists of a single function attribution and the accounts proposed by these philosophers leave much out of sight of what is achieved by functional explanations of the character of the item or behaviour to which the function is attributed. Function attributions are a first step in design and selection explanations and the etiological theories fail to account for the insights provided by the other steps. As I showed in chapter 6 one of the most interesting questions concerning functional explanations is: how can it be explanatory to compare real organisms with hypothetical organisms that may never have existed and may have played no role in the genesis of the trait to be explained. The etiological account ignores this question.

7.3.3 The normative role of function attributions

Both Millikan and Neander claim as one of the main merits of the etiological theory that it can account for the alleged fact that the parts and behaviours of organisms are classified on the basis of their function. In this section I argue that the assumption that biologists classify the items and behaviours of organisms in terms of their function is wrong.

Millikan starts the first chapter of her *Language, Thought and Other Biological Categories* (1984) by claiming that the different things which are called 'hearts' belong to the same category because they have the same function:

> That a heart is a heart certainly has something to do with pumping blood (Millikan 1984: 17)

However, as there are both diseased or malformed hearts that are unable to pump blood and things such as water pumps which are not hearts that are able to pump blood or that in fact pump blood (artificial hearts), the heart's function cannot be something the heart actually does or is capable of doing.

> It is not then the actual constitution, powers or dispositions of a thing that make it in a member of a biological category (Millikan 1984: 71)

Rather, what is relevant to being a heart is what it is “designed to” do or “supposed to” do. According to Millikan this notion of “designed to” or “supposed to” can be spelled out in a “naturalist, nonnormative and nonmysterious” way if ‘function’ is defined in historical terms:

> My claim will be that it is the “proper function” of a thing that puts it in a biological category, and that this has noting to do with its powers but with its history (Millikan 1984: 71).

Millikan repeats the idea that biological categories are historical categories in a paper published in 1986:

> Biological categories are carved out not by looking at the actual structure, actual dispositions, or actual functions of the organ or system that falls within the category but by looking at (or speculating about) its history (Millikan 1986: 51)

Hearts, for example, may have many different forms (they may be large or small, they may be three-chambered or four-chambered) and they may also be unable to pump blood, yet, they are all hearts. They are hearts, at least according to Millikan, because they were reproduced by mechanisms that proliferated in the course of the evolutionary history due to the fact that they pumped blood. In the same way the other parts of organisms are classified on the basis of what they did that caused them to be selected.

> Every body organ or system falls in the biological or physiological categories it does due to its historical connections with prior examples of kinds that have served certain functions or, typically, sets of functions. So whether or not it is itself capable of serving any of these functions, every organ or system is associated with a set of functions that are biologically “proper” to it, functions that helped account for the survival and proliferation of its ancestors. I call these functions “proper functions” of the organ or system (Millikan 1986: 52)

Millikan (1989b) talks of “function categories” instead of “biological categories” but the story remains the same: “function categories are *essentially* categories of things that need not fulfil their functions in order to have them” (p. 296, emphasis hers) and an analysis in terms of proper functions is the only one that is capable to account of this fact.

In the same vein, Neander (1991a: 180) claims that the etiological account of function is the only one that can account for “one of the most important theoretical roles of the notion [of function] in biology”, namely that “most biological categories are only definable in functional terms”. A definition of biological categories in terms of similarities in form does not work because biologists need categories that may include items that differ in form:

> Biologists need a category that ranges over different species, and hearts are morphological diverse: fish have a single pump with one auricle, but amphibians and most other [sic!][14] reptiles have a single pump

---

[14] Amphibians and reptiles are two different classes.

> with only one auricle, and while many reptiles have the ventricle partly partitioned, only crocodiles, birds and mammals have the two separate ventricles (Neander 1991a: 180).

Moreover, the relevant notion of function must be a notion not of what an item is capable of doing but of what an item is supposed to do:

> for the purpose of classifying hearts what matters is not whether the organ in question manages to pump blood, but whether that is what it is supposed to do. The heart that cannot perform its proper function (because it is atrophied, clogged, congenitally malformed, or sliced in two) is still a heart (Neander 1991a)

The etiological theory accounts for this ability by defining the proper function of an item in terms of its selection history. The definitions of ‘function’ provided by Cummins (1975), Wimsatt (1972) and Boorse (1976) , on the other hand, fail to account for this role of the notion of function in classifying the items and behaviours.

The arguments of Millikan and Neander can be represented in the following scheme:

(1) biological items and behaviours are categorized in terms of their function;
(2) a biological category may include items that fail to perform their function or are unable to perform their function;
(3) a biological category may include items that differ in form;
(4) the features ad (2) and (3) can be accounted for if the categories are seen as historical categories;
(5) it follows from (1) and (4) that function is to be defined in historical terms.

I argue that claim (1) is mistaken and that for that reason conclusion (5) does not follow, although (2), (3) and (4) are right.

Let me start by providing some counter examples against the claim that biological items are categorized in terms of their function. Perhaps all hearts have the function to pump blood, but there are many other biological categories that include items that differ in their function. A clear example is the category “tail”. The tails of sharks, bony fishes, tadpoles, crocodiles, beavers and whales have a propulsive function, the tails of cows and many other ruminants have the function to keep fly’s away, in many monkeys the tail has a function as a steering organ, some monkeys use the tail as a grasping organ as well, many mammals use their tail as a balancing organ, the hippopotamus uses its tail to spread its excrements and the tails of kangaroos have a supporting function.[15] Other examples are the categories ‘pharynx’, ‘wing’ and ‘limb’. In non-vertebrate chordates (such as tunicates and amphioxus) the pharynx has the function to filter food from water, in aquatic vertebrates the pharynx has a respiratory function. In most birds the

---

[15] Frogs, apes and man lack tails. The tail of birds is reduced. The function of the tail feathers that arise from the tail’s stump might have differs in different species. In many birds the tail feathers have a steering function during flight, the peacock’s tail has the function to attract females and to make them willing to mate.

wing has an important function in flight. Penguins are not able to fly and their wings have a function as a swimming organ. In most land vertebrates limb movements have a function in generating thrust. In salamanders, however, thrust is generated by means of undulations of the body and the limbs are stationary organs by means of which the push of the body movements is excerted on the ground. The forelimbs of lizards do have a function in generating thrust but their main function is in steering. The forelimbs of humans have a function in grasping objects but not in locomotion. The main function of the mole's forelimb is in digging.

My examples show that the categories biologists use to classify the parts of animal bodies do not only include items that differ in character (as Millikan and Neander emphasize) but also items that have different functions (as their first claim denies). I would agree with Millikan and Neander that many gross organs such as hearts, livers, kidneys and eyes have a similar function in many vertebrates, but this does not hold for such items as individual bones and muscles. I challenge the proponents of the etiological theory to come up with functional descriptions of categories like 'clavicle', 'humerus', 'thumb' or 'hand' which (i) applies to all the items that belong to this category in all the different classes of vertebrates, and (ii) is detailed enough to distinguish the items that belong to that category from all other parts of the body.

The fact that many biological categories include items with quite different functions shows that not all biological categories collect items in terms of their function. It does not follow, of course, that none does. Millikan and Neander might maintain that at least some of the categories used by biologists collect items in terms of their function, that those categories might include morphological diverse items, as well as items that do not perform their function or are not able to perform their function and that those features are best explained by assuming that those categories collect items in terms of their selected effects ("proper functions"). I argue against this view that there is a better explanation of those features, namely that those categories collect items on the basis of their evolutionary origin (homology).

All biology students learn in their first year[16] to distinguish between three ways to classify the items and behaviours of organisms, namely:

(a) on the basis of similarities in character (form and function$_1$),
(b) on the basis of their causal role (function$_2$),
(c) on the basis of common descent (homology).

All three types of criteria (note that 'selected effect' (fucntion$_4$) is not among them) have some use in biology and some categories are based on criteria of different types. For example, something is called a lung if it has a causal role in respiration (which is a criterion of the second type) and an invaginated structure (which is a criterion of the first type). However, the basic relation of *being the same thing* is homology. Two items in different individuals are (by defini-

---

[16] See for example: Keeton & Gould (1993: 521), Romer (1986: 9-11), Russell-Hunter (1968: 2)

tion) homologous (the same) if they are derived from a single item in a common ancestor. My heart is the same organ as your heart not because they are both supposed to pump blood but because they are derived from the same item of a common ancestor. Similarly, the lungs of mammals and the lungs of birds are the same organ not because they are both lungs, but because of their common evolutionary origin.

To avoid misunderstandings I should, perhaps, emphasize that my point is a point about categories not about names. Perhaps my heart is *called* a heart because it is supposed to pump blood, but my heart and your heart are the same organ not because they both are supposed to pump blood but because they have a common evolutionary origin. The names biologists use do not always reflect the categories. In many cases the same name is used for organs that belong to different categories. Insect hearts and vertebrates hearts are both *called* ‘hearts’ (perhaps because they are both supposed to pump blood), but as they are not homologous they belong to different categories. There are also many cases in which things with different names turned out to be homologous. The swim bladder of a teleost fish is a modified lung, despite the fact that lungs and swim bladders have different names, because the swim bladder originated as a lung in primitive osteichthyes (the common ancestors of both teleost fish and land vertebrates). The reason that the names do not reflect the categories is, of course, that many organs already had a name before their phylogenetic relationships were cleared up. If a category includes things with many different names biologists sometimes take resort to words which are seldom used in daily language. An example is the term ‘appendage’ which includes such things as fins, limbs, wings and flippers.[17]

The idea that the parts and organs of an organisms are to be classified primarily in terms of their function was abandoned by biologists in the beginning of the nineteenth century, long before Darwin. The term ‘homology’ was introduced by the English morphologist Richard Owen in 1843, but the idea that the parts and organs of members of different species are not only similar but in a certain sense “the same” goes back to the work of the French morphologist Étienne Geoffroy Saint-Hilaire (1772-1844). At the end of the eighteenth century the similarity between organisms belonging to different taxa was an important object of research in French morphology. The leading morphologist of that time, George Cuvier, maintained that those simi-

---

[17] For my purposes it is sufficient to define the notion of ‘homologous items’. Biologists also talk of ‘homologous characters’. The character of an item in certain individuals is homologous to the character of an item in other individuals if the items in question are homologous and the corresponding item of the most recent common ancestor had that character. The wings of bats and birds are homologous because they both originate from the forelimbs of a primitive terrestrial reptile-like ancestor. However, as the forelimbs of that common ancestor did not have a winged character, the winged character of the forelimbs of bats is not homologous to the winged character of the forelimbs of birds.

larities should be explained as the result of common functional requirements. Geoffroy argued against Cuvier that underneath the functional adaptations was a structural similarity that could not be explained by appeal to common functional requirements. He maintained that all animals are built in the same way and that the structural differences between organisms of different taxa are modifications of the same parts and organs in response to different functional requirements. Cuvier was quick to point out that the meanings of terms like ‘modification’ and ‘the same part or organ’ are unclear. Although Geoffroy’s work remained highly controversial he convinced the morphological community of his time that not all similarities between organisms could be explained on functional grounds. Two decades later, Owen (1843: 374-379) came up with a clear distinction between two ways to classify organs: homology and analogy. Owen defined “homologue” as “the same organ in different animals under every variety of form and function” and “analogue” as “a part or organ in one animal which has the same function as another part or organ in a different animal”. However, we had to wait until Darwin’s (1859) theory of common descent to clear up the vague ideas of ‘modification’ and of ‘the same organ’. According to this theory modification should be taken far more literally than was done before. Darwin himself did not delve into the notion of ‘sameness’, but he maintained that similarities between the same (i.e. homologous) item in members of different taxa can be explained as the result of common descent, and that the similarities between non-homologues items can be explained as the result of natural selection acting in the same direction. After Darwin, morphologists came to *define* ‘sameness’ (homology) in terms of a common origin. The first definitions of homology in terms of a common origin were given in 1870, by the English morphologist Lankester 1870[18] and, independently, by the German morphologist Gegenbaur.[19]

The definition of ‘sameness’ in terms of common descent (homology) is important in functional biology, first and foremost, because it enables biologists to say that a certain item has changed in character and/or has acquired a different causal role in the course of evolution. A famous example are the ossicles of the middle ear which were once part of the jaw of our fish-like ancestors and appear even earlier in the evolutionary history as part of the support of the gills of ancestral vertebrates. If in the course of the evolutionary history the character of a

---

[18] Actually, Lankaster proposed to drop the term ‘homology’ altogether, because that term was also used to indicate relations that had nothing to do with common origin such as the relation between the appendages of different segments of an arthopod. He proposed to call “structures which are genetically related, in so far as they have a single representative in a common ancestor” “homogenous” instead (Lankester 1870: 36).

[19] The first edition of Gegenbauer’s *Grundzüge der Vergleichende Anatomie* was published in 1859, the genetic definition of homology appears in the second edition of 1870, after Gegenbauer had adopted Darwin’s theory of common descent. In this latter edition ‘homology’ is defined as “the relation which obtains between two organs which have had a common origin, and which have also a common embryonic history” (Gegenbaur 1870: 64).

certain item changed extensively it might be difficult to track the evolutionary origin of that item. Two useful criteria are the position relative to other organs and the embryonic origin.

Furthermore, as both the character and the causal role of an item might change in the course of the evolutionary history, the concept of homology enables biologists to put items that differ in character and / or causal role in the same category.

The concept of homology enables biologists to say that a sperm cell which does not fertilize an egg cell is a sperm cell and that a malformed hearts is a heart. This is because (in contrast to what Millikan and Neander take for granted) being a sperm cell (that is being the same as the parts which in other organisms are called “sperm cells”) has nothing to do with fertilizing egg cells. Whether something is a sperm cell or not depends on its ancestry and has nothing to do with its function. Similarly a malformed heart is a heart not because of some function it is supposed to perform but because it originates from another heart. The case is not different from the swim bladder. A swim bladder is a modified lung because it is derived from an ancestral lung and a malformed heart is a modified heart because it is derived from an ancestral heart.

Millikan and Neander take it for granted that biologists categorize items in term of their function. Then they pose a riddle: how could this be since items that belong to a certain category need not perform their function and need not even be able to perform their function. The etiological theory answers this riddle by defining function in historical terms. My answer has been that Millikan and Neander are mistaken in their initial assumption. Biologists classify items primarily in terms of homology. The riddle does not arise. Malformed hearts are hearts because they are homologous to other hearts, not because of their function.

At this point someone might want to interrupt and say “wait a moment, perhaps your appeal to homology shows how malformed hearts can be hearts, but you fail to account for the fact that hearts that are not able to pump blood still have the function to pump blood. In other words how do you account for the biologist’s ability to attribute functions to items that do not perform their function?”. My answer is consist of two parts. First, on my theory function attributions are generalizations about the causal role of the members of a certain homology-category in a certain group of organisms. Vertebrate hearts have the causal role to pump blood because this is how hearts contribute to the maintenance of the organism in those vertebrates that are able to circulate blood. Second, one should not exaggerate the importance of attributing functions to items that are not able to perform their function. The point of a function attribution is to tell us how a certain item or behaviour fits into the organism, not to tell us what malformed items should do.

Both Millikan and Neander tend to distinguish between ‘having a function’ and ‘serving a function’. Statements of the form ‘the function of ... is ....’ are used to express the functions an item or trait has, statements of the form ‘.... functions as .....’ are used to express the functions an item or trait serves. For example, Millikan says:

> the definition of "proper function" is intended to explain what it is for an item to *have* a function or purpose, but not what it is for an item to function *as* something. Robert Cummins (1980, 1984) has given us a definition of function that is probably best construed as a theoretical definition [...] of "function as" in some contexts (Millikan 1989b: 293).

She adds:

> A diseased heart may not be capable of pumping, of functioning *as* a pump, although it is clearly its function, its biological purpose, *to* pump, and a mating display may fail to attract a mate although it is called a "mating display" because its biological purpose is to attract a mate (Millikan 1989b: 294).

Similarly, Neander maintains that

> There is a distinction between a trait *serving* a function and *having* a function [...] it is the function of all kidneys, diseased or otherwise, to filter the blood, but if they are very diseased they will be incapable of performing, or serving, this function (Neander 1991b: 465).

Such utterances betray a misunderstanding of the role of function attributions in functional biology. The point of a function attribution is not to tell us what a (possibly malformed) item should do,[20] but to tell us how a certain item fits into the organism. On the etiological theory the notion of 'serving a certain function' is dependent on the notion of 'having a certain function'. If a certain item has a selection history it has a function and it may perform that function or it may fail to perform it.[21] The functions which an item has are in a certain sense intrinsic to that item and this is why they are called *proper* functions.[22] However, in functional biology it is the other way round. A function is not something intrinsic to an item but a position in a system that performs a certain task. Function predicates are position-predicates and function attributions situate an item or behaviour into a system by telling us how that item or behaviour contributes to that system's tasks by listing the subtasks performed by that item or behaviour. Statements which say that a certain item has a certain function or which say what the functions of a certain item are, are generalizations about the causal roles which the members of a certain category of homologous items perform in a certain group of animals. This means that to say that a certain item is not able to perform a certain function it has, is just another way of saying that that item is not able to perform the causal role performed by homologues items of other

---

[20] Neither are biologists interested in distinguishing functional and accidental *effects* as for instance Larry Wright (1973, Wright 1976) maintains. A function is a position-predicate

[21] Neander (1991b: 465) says that "a trait need not have a function in order to serve it", but it remains unclear how on the etiological theory one can say that an item serves a function if that item does not have that function.

[22] Millikan (1993a: 31) claims that she has invented the combination "proper function" and that by adding the word "proper" in front of "function" she wanted to "contrast a thing's 'proper' or *own* function with functions imposed on it or accidental to it" (note 1).

animals in the same group. Perhaps there is another item that performs that role or perhaps the position remains empty as a result of which the organism is probably not viable.

Millikan and Neander suggest that function attributions have two normative roles, namely to put items into a cross species category and to tell us what an item should do. I have argued that in functional biology the first role is performed by the notion of 'homology' and that the second role is unimportant in functional biology. I have also argued that the point of a function attribution is not to tell us what a (possible malformed) item should do but to situate an item in a system that performs a certain task. Millikan and Neander put the chart before the horse by defining 'performing a function' in terms of 'serving a function'.

7.3.4 Conclusion

On the etiological account function attributions are normative and teleological. They tell us what an item should do or what it is supposed to do and they do so on the basis of what that item is for (that is why it was brought about—why it was selected). I have argued that this normative and teleological concept of function does not apply to the term 'function' as this term is used functional biology. The point of function attribution in functional biology is to position an item in a system, not to say what it should do.

That the etiological theory does not apply to talk of functions in functional biology is shown by the following considerations. First, when supporting function attributions biologists do not provide the kind of evidence that is needed to draw conclusions about why a certain item, behaviour or trait was selected in the past. Second, the way in which biologists appeal to functions in explanations is quite different from the way in which such appeals should be used according to the etiological theory. On the etiological account function attributions are intrinsically explanatory in that they need not be combined with other statement to tell us why a certain trait is present in the place where it is present. Biologists use function attribution in capacity, design and selection explanations. Both design and selection explanations explain in a certain sense why a trait is present. In both kind of explanations start by attributing a function to the item concerned. This function attribution is the first step of the explanation, not the complete explanation. In the case of a design explanation the next part says that the character which the item has is useful because that function is better performed if the item is the way it is than in some other conceivable way. In the case of an evolutionary explanation the next part says that the character which the item has evolved because variants having that character performed that function better than competing variants. In both cases the function attribution is not intrinsically explanatory but yields an explanation only when combined with another statement. The function attribution is not dependent on this second statement. Third, the claim of Millikan and Neander that functional biologists need a normative notion of function in order to create interspecies categories that unite items that differ in form and that may contain items that do not perform their function or are not able to perform their function turned out to be untrue. Biologists may use the notion of ‘homology’ to create such interspecies categories.

I draw the conclusion that the notion of proper function as this notion was invented by Neander and Millikan does not apply to the use of the term ‘function’ in functional biology and is if no help to understand the explanatory practice in that discipline.

Chapter 8: A non-causal account of design explanation

8.1 Introduction

Let us take stock. As I explained in chapter 1, I aim to understand a certain practice in functional biology. Functional biology is that part of biology which is concerned with the life of individual organisms: the way they are built, the way they work and the way they behave. As I have shown by many examples, functional biologists often explain why an organism is built the way it is built, why it works the way it works, or why it behaves the way it does, by appeal to the causal roles of the items and behaviours in the maintenance of the organism and the survival value of the character those items and behaviours have. I aim to understand what this activity adds to our knowledge. This activity is usually called ‘functional explanation’ by those who engage in it. I use the term ‘design explanation’ to avoid confusion with other uses of the term ‘functional explanation’.

In chapter 5 I argued that Cummins gives a satisfactory account of the meaning of the notion of function as causal role, but not of the use of attributions of function as causal role in design and evolutionary explanations. In chapter 2 and in chapter 5 I distinguished the notion of function as causal role from the notion of function as survival value and I presented and defended a new account of the latter notion.

In chapter 4–7 I discussed the different accounts of functional explanation presented by philosophers. I argued that those philosophers are (in part) concerned with different kinds of explanations. There are at least three different kinds of explanations which are called functional explanations (at least by some philosophers):

1) functional explanation$_{1:}$ functional analyses in Cummins’s sense (capacity explanations)
2) functional explanation$_2$: the kind of explanations in which I am interested (design explanations),
3) functional explanation$_3$: functional explanations in the sense of the etiological theory, (selection explanations).

I argued that the different accounts of functional explanation presented by philosophers up to now fail to give a satisfactory account of design explanations.

In this chapter I detail my own account. The basic idea of my account is that design explanations relate the way an organism is built, the activities of its parts, its behaviour and the condition of the environment in which it lives in terms of what is needed or useful to survive and reproduce rather than in terms of causes. In section 8.2 I describe the structure of design explanations. I argued that design explanations typically start with an attribution of a causal role and then proceed in one of the following two ways. If they address a type (4a) question (why is it useful to have a certain item or perform a certain behaviour?) they continue by explaining why it is useful to perform that causal role by pointing to the demands imposed on the organism in the conditions that apply to the organisms in study. If they answer a type (4b) (why does a certain item or behaviour has the character it has?) they continue by explaining why it is useful to perform that causal role in the way in which it is performed by pointing to the demands upon that causal role in the conditions in which the organisms in study live. In the next two sections I am concerned with the subject of design explanations. In section 8.3 I discuss the different kind of relations that are the subject of the different kinds of explanation in functional biology. I contend that explanations show us how a certain phenomenon fits into the structure of the world and I argue that there are three kinds of fundamental relations in the organic world: causal relations at the individual level, causal relations at the population level and functional interdependencies at the individual level. In section 8.4, I answer the question what design explanations add to our knowledge. My answer is to that question basically this: design shows us how the trait to be explained fits into the structure of functional interdependencies of a living organisms.

8.2 The structure of design explanation

8.2.1 The core of a design explanation

Design explanations are answers to a question of the following form:

‘why do *s*-organisms have / perform t_1 rather than t_2, t_3, … t_n?’.

In which:

s	a set of organisms[1] (the phrase ‘s-organisms’ refers to the members of that set)
t_1	a trait (i.e. the presence or character of a certain item or behaviour) of *s*-organisms
t_2, t_3, … t_n	traits which *s*-organisms lack

I shall call the set {t_1, t_2, t_3, … t_n} the ‘reference class’.[2] Trait t_1 will be called ‘the trait in question’, t_2, t_3, … t_n will be called ‘the alternative traits’.

---

[1] Note that *s* does not necessarily consists of organisms of one taxon. The set might be taxonomically heterogeneous. That is, it may lump together organisms that belong to more than one taxonomic group (e.g. fishes) and/or it may excluded certain members of a certain group (e.g. green plants, land mammals).

[2] Note that my notion of reference class is akin to but different from Van Fraassen’s (1980) notion of ‘contrast class’. The contrast class in Van Fraassen’s sense would be: {‘s-organisms have / perform t_1’, ‘s-organisms have / perform t_2’, ‘s-organisms have / perform t_3’, ..., ‘s-organisms have /perform t_n’}. The reason why I use my

In the preceding sections I have discussed many examples of such questions. Among others:

- why do snakes have a forked tongue rather than a blunt one?
- why do "larger" organisms have rather than lack a circulatory system?
- why do fishes respire by means of gills rather than by means of lungs?
- why do mountain gazelles mark their territories by dung piles rather than by scent marks?
- why do birds remove the empty egg shell after the chick has hatched rather than leave it where it is?

However, not every why-question of this form counts as a request for a design explanation. Why-questions in biology are notoriously ambiguous. In section 2.3.2 I distinguished three types of why-questions:

1) questions that ask for causes at the level of an individual organism, such as the mechanisms that bring about a snakes forked tongue or the signals that release the egg shell removal behaviour in birds (why$_1$-questions),
2) questions that ask for the utility of a trait, such as the utility of a forked tongue or the utility of the egg shell removal behaviour (why$_2$-questions),
3) questions that ask for evolutionary causes, such as the mechanisms that brought about the forked tongue or the egg shell removal behaviour in the course of the evolution (why$_3$-questions).

Design explanations are answers to why$_2$-questions. So the general form of a request for a design explanation is:

'why$_2$ do *s*-organisms have / perform t_1 rather than t_2, t_3, … t_n?'.

The core of an answer to a why$_2$-question consists of two statements of the following form:

(1) *s*-organisms live in condition c_u.

(2) in condition c_u trait t_1 is more useful than trait t_2, t_3, … t_n.

In which:

c_u a conjunction of one or more conditions of organisms and/or the environment in which they live[3]

For example, the core of the answer to the question 'why$_2$ do fishes respire by means of gills rather than by means of lungs?' consists of the following statements:

---

own notion rather that Van Fraassen's is that I want to say such things as 'the trait in question is the best one in the reference class'.

[3] The subscript *u* is added to the *c* of 'condition' to emphasize that *c* refers to the conditions in which a certain trait is useful not to the conditions that bring about (cause) that trait.

(1) Fishes typically live in water.
(2) In water it is more useful to respire by means of gills than by means of lungs.

I will represent the train of thought of design explanations (functional explanations$_2$) in the following form:

(1) *s*-organisms live in condition c_u.
(2) in condition c_u trait t_1 is more useful than trait t_2, t_3, … t_n.
----
(3) that's why$_2$ *s*-organisms have / perform t_1 rather than t_2, t_3, … t_n.

For example:

(1) Fishes typically live in water.
(2) In water it is more useful to respire by means of gills than by means of lungs.
----
(3) That's why$_2$ fishes respire by means of gills rather than by means of lungs.

The final statement in this series indicates (i) what question is addressed by the explanation, and (ii) that at this point of the reasoning that question is supposed to be sufficiently answered. It is worth emphasizing that this train of thought is not an argument in the logical sense of 'argument': it does not consist of a series of premises that support a conclusion. It is not required that the part after 'that's why' in the final statement follows logically from the premises. In order to be a good explanation it is required that

- the explaining statements are well supported,
- any other requirements concerning the scientific status of the explaining statements are met,
- the statements in the scheme are related in the way indicated by the symbols in the scheme,
- any other requirements concerning the (inferential) relation between the different statements in the scheme are met,
- the symbols in the scheme are filled in in accordance with the filling instructions.

The first statement in this train of thought is an empirical generalization. The second is an attribution of survival value, which is a lawlike statement. As I will explain in section 8.3.4, this statement states that condition c_u is a non-redundant part of a condition that is sufficient for the utility of t_1 over t_2, t_3, … t_n. In such cases it is said that c_u imposes a demand for t_1 on the relevant item or organism or that c_u demands for t_1. For example, living in water imposes a demand for gills on fishes and using the tongue in chemical tropotaxis imposes a demand for being forked on the snakes tongue. Usually t_1, t_2, t_3, … t_n can be viewed as different states of the same character or different determinations of the same determinable property.

There are several ways in which an item or behaviour can be useful. Roughly speaking, three criteria for usefulness can be distinguished: (1) the efficiency with which a certain task is performed, (2) the effects on survival, and (3) reproductive success. The first criterion is, for example, used in Schwenk's explanation of the snake's forked tongue. Schwenk observes that a forked tongue is more useful than a blunt one because a blunt one would be of no use in the mechanism of chemical tropotaxis. The second criterion is used in the explanation of why fishes have gills rather than lungs. This explanation points out that organisms that live in water are less viable if they have lungs than if they have gills. Similarly the explanation of why e.g. Vertebrates transport oxygen by means of a circulatory system points out that organisms that are larger than a certain size are more viable if they transport oxygen than if they rely on diffusion alone. The third criterion is employed in Tinbergen's explanation of the egg shell removal behaviour. Tinbergen points out that birds that are predated by carrion crows and herring gulls have more reproductive success if they remove the empty egg shell than if they leave it near the nest.

As I showed in the preceding chapters, philosophers have different opinions about what the hallmark is of functional explanation. Most of the classical authors characterize functional explanations (usually they called them 'teleological explanations') as explanations that appeal to effects rather than to causes. Cummins (1975, 1983) has defended the view that functional explanation (he calls them 'functional analyses') are characterized by the fact that they appeal to causal roles in maintaining a certain capacity. On the etiological view the hallmark of functional explanations is their teleological nature, that is the alleged fact that they appeal to what a certain item evolved for.

In my view the analyses of those other philosophers apply in part to different kinds of explanations (different from each other and from the kind of explanations that are central to this thesis). Cummins's analysis applies to capacity explanations. In chapter 7 I have raised doubts the kind of explanations envisaged by the teleological theory, but anyway such explanations are different in kind from both capacity explanation and design explanations. The hallmark of design explanations is their appeal to utilities. As utilities are usually assessed in relation to causal roles (function$_2$), design explanations typically start with a function$_2$ attribution (attribution of a causal role) but it is the appeal to utilities and not the appeal to causal roles that gives design explanations (functional explanations$_2$) their special status and what makes them different from causal explanations.

8.2.2 Appeal to causal roles

In the previous section I emphasized that the hallmark of design explanations (functional explanations$_{2)}$ is their concern with the utility of a certain trait. Nevertheless attributions of causal role (function$_2$ attributions) play an important role in design explanations. This is because utilities are usually assessed in relation to causal roles. The role of attributions of causal roles in design explanations is the subject of this section.

Roughly speaking, two types of design explanations (functional explanations$_2$) might be distinguished (see section 2.3.2, and the examples of section 6.2): design explanations that answer a type (4a) question (why is it useful to perform a certain causal role?) and design explanations that answer a type (4b) question (why is it useful that a certain item or behaviour has a certain character?). Explanations of both types typically start with an attribution of a causal role *f* (say respiration) to an item or behavioural pattern *i* (such as the gills of fishes). From there they proceed in different ways. The first type of design explanation is concerned with the survival value of the attributed causal role. It addresses the question 'why$_2$ do s-organisms have an item / behaviour that performs task *f*?' (e.g. 'why$_2$ do fishes have a respiratory system?'). The second type of design explanation is concerned with the specific way in which the attributed causal role is performed. It addresses the question 'why$_2$ is task *f* performed in the way it is performed rather than in some other way' (e.g. 'why$_2$ do fishes respire by means of gills rather than lungs?').

An example of a design explanation of the first type (which answers a type 4a question) is Krogh's (1941) explanation of the need for a system that transports oxygen, discussed in section 4.2.3. This explanation can be reconstructed as follows:[4]

(1) The distance between the inner organs and the periphery of vertebrates is more than 1 mm.
(2) If the distance between the inner organs and the periphery of an organism is more than 1 mm it is useful to that organism to transport oxygen.
----
(3) That's why$_2$ vertebrates have a system that transports oxygen.

More generally, the core of a design explanation (functional explanation$_2$) which explain why an organism has an item that performs a certain task has the following scheme:

(1) s-organisms live in condition c_u.
(2) In condition c_u it is useful to perform *f*.
----
(3) That's why$_2$ s-organisms have an item / behaviour that performs *f*.

---

[4] This a *very* rough sketch of Krogh's explanation. Recall that Krogh's main achievement was the application of Fick's law of diffusion to the problem of circulation and respiration. This law is not even mentioned in this rough sketch. This is because this section is concerned with the role of attributions of causal roles in survival value explanations. The explanatory role of physical laws is discussed in section 8.2.3. In this section I give a more detailed sketch of Krogh's explanation.

In which:

f a causal role (function$_2$).[5]

The first statement in this scheme is an empirical generalization about the conditions in which the organisms live to which the explanation applies. The second statement is an attribution of survival value. The phrase 'in condition c_u it is useful to perform *f*' should be read as 'in condition c_u it is more useful to be able to perform *f* than to live without that capacity'. This is a lawlike statement.

An explanation with this kind of core is often given in answer to a question of the following type:

why$_2$ do *s*-organisms have item / behaviour *i*?

The answer to this question has the following structure:

(1) *s*-organisms live in condition c_u.

(2) in condition c_u it is useful to perform *f*.

(3) in *s*-organisms item / behaviour *i* contributes to task *f*.

----

(4) That's why$_2$ *s*-organisms have / perform *i*.

In which:

i a certain type of item or behaviour[6] (the phrase 'item / behaviour *i*' refers to items or behaviours of type *i*).

For example:

(1) The distance between the inner organs and the periphery of vertebrates is more than 1 mm.

(2) If the distance between the inner organs and the periphery of an organism is more than 1 mm it is useful to that organism to transport oxygen.

(3) In vertebrates the circulatory system transports oxygen.

----

(4) That's why$_2$ vertebrates have a circulatory system.

The first statement in this scheme is an empirical generalization about the conditions in which the organisms to which the explanation applies live. The second statement (an appeal to survival value) is a lawlike statement saying that in the conditions stated in the first statement it is useful to perform a certain task (function$_2$). The third statement (an attribution of a causal role

---

[5] As I discussed in section 5.1 I agree with Cummins (1975) that functions$_2$ (causal roles) are singled out by their role in capacity explanations.

[6] Recall that items are grouped together on the basis of homology (chapter 7). If items / behaviours were identified in terms of their role the first statement in this scheme would be tautologous.

(function$_2$)) is an empirical generalization about how that task is performed in the organisms to which the explanation applies. It says that in those organisms that causal role is performed by the item or behaviour to be explained.

This first kind of design explanation is concerned with the need to perform a certain causal role, it says nothing about the character of the item or behaviour that performs that causal role. The second kind of design explanation is concerned with the way in which a certain causal role is performed and, therefore, with the character of the item or behaviour that performs the causal role. The core of this second type of design explanation (which answers a type 4b question) has the following structure:

(1) *s*-organisms live in condition c_u.

(2) In condition c_u it is more useful to perform *f* by means of an item / behaviour that has character s_1 than by means of an item / behaviour that has character s_2.

----

(3) that's why *s*-organisms perform *f* by means of an item / behaviour that has character s_1 rather than by means of an item / behaviour that has character s_2.

In which:

s_1 a description of *i*'s form or activity (function$_1$),

s_2 a description of a form or activity that does not apply to *i*.

An example is the explanation of why fishes respire by means of gills rather than lungs given in section 6.2.2. This explanation can be summarized as follows:[7]

(1) Fishes live in water.

(2) In water it is more useful to respire by means of invaginated structures (gills) than by means of evaginated structures (lungs).

----

(3) That's why$_2$ fishes use invaginated structures (gills) rather than evaginated ones (lungs) to respire.

Another example is the explanation of Habibi c.s. (1993) of why mountain gazelle's mark their territories by dung piles rather than by scent marks (example 3.3. of section 2.2.3). This explanation can be reconstructed as follows:[8]

---

[7] Just as in the scheme of Krogh's presentation this is a very rough sketch. I have left out the elaboration on why gills are more useful than lungs. This part of the explanation is discussed in section 8.2.3

[8] This is, once again, a very rough sketch. I have left out Habibi c.s. explanation of why it is the case that in large territories it is more useful to mark territories by means of dung piles rather than scent marks. This part of the explanation is discussed in section 8.2.3.

(1) Mountain gazelle's have large territories.
(2) In case of large territories it is more useful to mark territories by means of dung piles than by means of scent marks.
----
(3) That's why$_2$ mountain gazelle's mark their territories by dung piles rather than by scent marks.

In morphology, this kind of explanation is often brought up in answer to a question of the following form:

why$_2$ does item *i* of *s*-organisms has character s_1 rather than s_2?

The answer to such a question has the following structure:

(1) In s-organisms: item *i* performs causal role *f*
(2) s-organisms live in condition c_u
(3) In condition c_u it is more useful to perform *f* by means of an *i* that has character s_1 than by means of an *i* that has character s_2.
----
(4) That's why$_2$ in *s*-organisms item *i* has structure s_1 rather than s_2

A clear example is Schwenk's (1994) explanation of why snakes have a forked tongue (example 2.3 of section 2.2.2):

(1) The tongue of snakes has a causal role in trail following
(2) Snakes follow trails by comparing chemical stimuli simultaneously sampled at two sides
(3) In order to sample chemical stimuli simultaneously at two sides it is more useful to have a forked tongue than a blunt one
----
(4) That's why snakes have a forked tongue rather than a blunt one.

The favourite explanation of many philosophers also follows this scheme:

(1) The function$_2$ (causal role) of hearts of vertebrates is to propagate the blood.
(3) In order to propagate the blood it is more useful if the heart beats than if it stands still.
----
(4) That's why the heart of vertebrates beats.

Note that in this example (just as in the others) the function$_2$ attribution (attribution of a causal role) is a first step in the explanation, not the complete explanation. This in contrast to popular beliefs.

In ethology the second kind of explanation is often invoked to answer a question of the following form

why$_2$ do s-organisms perform behaviour s_1 rather than s_2?

in which:

s_1 describes a behaviour s-organisms perform,

s_2 describes a behaviour s-organisms do not perform.

The answer to such a question has the following form:

(1) In s-organisms: behaviour s_1 performs causal role *f*.

(2) s-organisms live in condition c_u.

(3) In condition c_u it is more useful to perform *f* by means of a behaviour that has character s_1 than by means of a behaviour that has character s_2.

----

(4) That's why$_2$ s-organisms perform s_1 rather than s_2.

In this section I have shown that function attributions have two important explanatory roles in design explanations (functional explanations$_2$). Utilities are usually assessed in relation to functions$_2$ (causal roles). For that reason, design explanations typically contain a statement which says that it is more useful to perform a certain causal role (function$_2$) in the way it is performed than in some other way. Furthermore, because utilities are usually assessed in relation to functions$_2$ (causal roles) it is a good strategy to start to look for a function$_2$ (causal role) of an item or behaviour if you want to explain the presence or the specific character of that item or behaviour by means of a design explanation. Hence, design explanations of the presence or the character of an item or behaviour typically start with a function$_2$ attribution (attribution of a causal role). They then proceed to show that it is more useful to perform that function by an item or behaviour with the character which that item or behaviour actually has than by an item or behaviour with some other character. Function$_2$ attributions (attributions of a causal role) are the first part of such explanations, but not the complete explanation.

8.2.3 Optimality claims and requirement claims

In section 8.2.1 I stated that the core of a design explanation consists of two claims. One states that in certain conditions the traits to be explained are more useful to the organisms that have them than some other traits which they have not. The other claims that those conditions apply to the organisms concerned. In section 8.2.2 I added that utilities are usually assessed in relation to functions$_2$ (causal roles). The utility claim, therefore, usually takes the form of a claim about the utility of performing a certain function$_2$ (causal role) in a certain way. Claims about utilities may vary in strength. I distinguish two types: optimality claims and requirement claims. Optimality claims claim that the trait in question is the best one in the reference class. They have the following form:

in condition c_u trait t_1 is more useful than each of the following traits: t_2, t_3, … t_n

For example: ‘if the tongue is used in trail following by means of chemosensory tropotaxis a more deeply forked tongue is more useful than a less deeply forked tongue’. The ultimate criterion for utility in optimality claims is inclusive fitness. That is, ultimately it is claimed that the inclusive fitness of the organisms in question is higher if they have trait t_1 than if they have one of the other traits in the reference class instead of t_1. However, as utility is usually assessed in relation to function$_2$ (causal role) one often finds claims to the effect that the trait in question performs a certain function$_2$ more efficiently than any other trait in the reference class (for example ‘a more deeply forked tongue is more efficient in performing chemosensory tropotaxis than a less deeply forked one’). In such cases it is tacitly assumed that efficient performance results in greater fitness.

Requirement claims claim that the trait in question is the only one in the reference class that works. They have the following form:

> in condition c_u trait t_1 is the only one that is useful among the following traits: t_1, t_2, t_3, … t_n

For example: ‘if an organism lives on land, lungs are useful but gills not’. I discuss the precise meaning of requirement claims and related notions in section 8.4. Depending on the criterion for utility one may distinguish between viability claims and functionality claims. Viability claims claim that the trait in question is the only one in the reference class that enables the organism to survive and reproduce. Functionality claims claim that the trait in question is the only one in the reference class that enables the organism to perform a certain function$_2$ (causal role). The distinction between viability claims and functionality claims is a gradual one. The above claim about the utility of lungs is clearly a viability claim because it implies that land organisms that respire by means of gills are not viable (cannot get enough oxygen to survive and reproduce). Since biologists are usually interested in the survival value of a certain trait, real examples of pure functionality claims are rare. Schwenk’s claim that a forked tongue is useful in chemosensory tropotaxis but a blunt one not has for example both functionality and viability aspects. It is formulated as a claim about what is required to perform a causal role, but it has implications about survival: Schwenk points out that the ability to perform chemosensory tropotaxis is important in finding preys and mates.

If the reference class is in a certain sense complete (e.g. because the reference class contains all possible states of a certain character or because the presence of a certain item is compared to its absence) one says that the trait in question is *needed* (in order) to survive and reproduce or (in order) to perform a certain function. For example, Krogh claimed that larger organisms need a system of convection in addition to diffusion. The explanation of why fishes have gills rather than lungs points out that fishes need gills (rather than lungs) to survive and reproduce. Schwenk points out that snakes need a forked tong to perform chemosensory tropotaxis.

Simple design explanations might be classified either as optimality explanations or as requirement explanations depending on whether the utility claim is a optimality claim or a requirement claim. However, many design explanations are complex in the sense that they contain more than one utility claim and in that case they might contain both optimality claims and requirement claims.

8.2.4 Requirements, problems, (dis)advantages

As I discussed in the previous sections, the core of a design explanation consists of two claims. One states that in certain conditions the traits in question are more useful to the organisms that have them than the alternate traits. The other claims that those conditions apply to the organisms concerned. Utilities are usually assessed in relation to functions$_2$ (causal roles). The utility claim, therefore, often takes the form of a claim about the utility of performing a certain function$_2$ (causal role) in a certain way. In the previous section I stated that claims about utilities may vary in strength. They may claim that the trait in question is the only one in the reference class that works (requirement claims) or they may claim that the trait in question is the best one in the reference class (optimality claims). Most design explanations do not merely claim *that* the trait in question is more useful than the other traits in the reference class, they also make plain *why* this is the case. This elaboration is the subject of the present section. I distinguish between requirement explanations and optimality explanations.

Requirement explanations

Requirement explanations explain why a certain trait is needed by identifying a problem into which an organism would run if it lacked the trait in question. The problem consists in a certain requirement not being met. For example, Krogh points out that the inner cells of larger organisms would not get enough oxygen if those organism relied on diffusion alone. Note that the problem is not a problem that is experienced by the real organism. It is a problem that would arise if the trait in question in a real organism was replaced by one of the alternative traits. In other words, the problem is identified by comparison with a *hypothetical* organism. Claims about such hypothetical organisms are established by experiments or calculations. These experiments and calculations aim to show (1) that and (2) why such a hypothetical organism would not be able to survive and reproduce / to perform a certain causal role, and (3) that and how the addition of the trait to be explained would solve this problem. The result of this analysis is expressed in a functional counterfactual of the form “if s-organisms had one of the alternative traits instead of the trait in question they would have problem *P*”.

An example of a viability explanation along these lines is the explanation of why vertebrates have (rather than lack) a circulatory system (suggested by Krogh):

(1) the distance between the inner organs and the periphery of vertebrates is more than 1 mm.

(2) for all organisms: if the distance between the inner organs and the periphery is more than 1 mm they need a system of convection in addition to diffusion.

(3a) in Vertebrates the circulatory system provides a system of convection in addition to diffusion.

(3b) if Vertebrates lacked a circulatory system they would not have a system of convection.

----

(4) that's why Vertebrates have a circulatory system.

The first statements specifies certain conditions in which vertebrates live. The second states a requirement which applies in these conditions. The third and the fourth stated that this requirement is met if the trait in question is present but not if it is absent.

More generally, viability explanations have the following structure:

(1) s-organisms live in condition c_u.

(2) in condition c_u organisms can survive and reproduce only if requirement *r* is met.

(3a) if s-organisms have / perform t_1 requirement *r* is met.

(3b) if s-organisms have / perform t_2 requirement *r* is not met.

……

(3n) if s-organisms have / perform t_n requirement *r* is not met.

----

(4) that's why_2 *s*-organisms have / perform t_1 rather than t_2, t_3, ..., or t_n.

Schwenk's (1994) explanation of why snakes have a forked tongue is an example of an explanation that appeals to what is required to perform a causal role:

(1a) Snakes and lizard use their tongue to sample chemicals

(1b) Snakes and lizards follow trails by comparing chemical stimuli simultaneously sampled at two points

(3) In order to compare chemical stimuli simultaneously sampled at two points an organism must be able to sample chemical stimuli simultaneously at two points

(4) In snakes and lizards this requirement is met if the tongue is forked but not if it is blunt

----

(5) That's why the tongues of snakes and lizards are forked rather than blunt

The first two statements are function_2 attributions (attributions of causal roles). I have labelled them (1a) and (1b) to emphasize that they have the same character. The third one states a requirement. One might call this kind of statement an 'attribution of a need'. In this case the need / requirement is not conditional, so there is no need for a specification of the conditions in which snakes and lizards live. I have labelled the requirement claim (3) because this facilitates comparison with the general structure (to be presented shortly) which provides for a specification of the conditions in which the organisms in question live. The fourth statement says that this requirement is met if the organism has the trait in question but not if it has one of the alternative traits.

The general structure of a design explanation of the structure of an item by appeal to what is required to perform a certain function$_2$ (causal role) is:

(1) In s-organisms: item i performs causal role (function$_2$) f.
(2) s-organisms live in condition c_u.
(3) in condition c_u: causal role f is physical possible only if requirement r is met.
(4a) in s-organisms: if item i has character s_1 then requirement r is met.
(4b) in s-organisms: if item i has character s_2 then requirement r is not met.

……

(4c) in s-organisms: if item i has character s_n then requirement r is not met.

----

(5) That's why$_2$ in s-organisms item i has structure s_1 rather than s_2, s_3, ..., or s_n.

Requirement explanations often point out a problem without mentioning the relevant requirement explicitly. In that case one gets something like:

(1) s-organisms live in condition c_u.
(2a) in condition c_u: if organisms had / performed t_2, t_3, ..., or t_n instead of t_1 problem p would occur.
(2b) in s-organisms: if t_1 is present problem p does not occur.

----

(3) that's why$_2$ s-organisms have / perform t_1 rather than t_2, t_3, ..., or t_n.

For example:

(1) In vertebrates the distance between the inner organs and the periphery is more than 1 mm.
(2a) If the distance between the inner organs and the periphery is more than 1 mm vertebrates would not be able to meet their demand for oxygen if they had to rely on diffusion alone.
(2b) the presence of a system of convection would solve this problem.
(3) in Vertebrates the circulatory system provides a system of convection in addition to diffusion.

----

(4) that's why Vertebrates have (rather than lack) a circulatory system.

Actually, as indicated by the numbers, this example is more complex than the abstract scheme says. It starts with stating the conditions that apply to the organisms in question (vertebrates) (1). The second statement (2a) points to a problem these organisms would have if they lacked the trait in question (a circulatory system). The third statement (2b) points out a requirement in the form of a general solution to this problem (namely having a system of convection in addition to diffusion). The fourth statement (3) states how this general solution is implemented in vertebrates.

In sum: requirement explanations point out a problem the organisms in question would have in the conditions in which they live if the trait in question were replaced by one of the alternative traits. A problem means that a certain requirement is not met. The ultimate requirement is that the organism is viable and productive, but in many cases the explanation mentions a requirement which is derived from this ultimate one.

Optimality explanations

Optimality explanations proceed in a similar way. They point out that the organisms in question *would* have certain disadvantages in the conditions in which they live if the trait in question *were* replaced by an alternative. The ultimate disadvantage is that their fitness would be lower if the organism had an alternative instead of the trait in question. In many cases the explanation employs another criterion such as the efficiency by which a certain function$_2$ (causal role) is performed. This criterion is supposed to correlate with fitness.

An example of such an explanation is the explanation of Tinbergen and his students (1962) of why black-headed gulls remove the empty egg shell after the chick has hatched. This example was discussed in section 2.2.3 (example 3.1). This example can be represented as follows:

(1) The eggs of black-headed gulls are subject to predation by herring gulls and carrion crows.
(2) In the condition that eggs are predated by herring gulls and carrion crows it is useful to minimize the risk that the nest is discovered by herring gulls and carrion crows.
(3) The nest is less easily discovered by herring gulls and carrion crows if the empty shells are removed than if they are left.
----
(4) That's why$_2$ black-headed gulls remove the empty egg shell rather than leave it near the nest.

The first sentence states a condition in which the organisms in question live. The second states a utility criterion which applies in these conditions. The third sentence states that on this criterion the trait in question is more useful than the alternative traits.

Optimality explanations often point out disadvantages directly, without mentioning an explicit utility criterion. In that case the explanation has the following structure:

(1) s-organisms live in condition c_u.

(2a) in condition c_u if organisms had / performed t_2 instead of t_1 problem p_1 would occur.

(2b) in condition c_u if organisms had / performed t_3 instead of t_1 problem p_2 would occur.

....

(2n) in condition c_u if organisms had / performed t_n instead of t_1 problem p_o would occur.

(2o) in s-organisms: if t_1 is present problem p_1, ..., p_o do not occur.

----

(3) that's why$_2$ *s*-organisms have / perform t_1 rather than t_2, t_3, ..., or t_n.

For example:

(1) The eggs of black-headed gulls are subject to predation by herring gulls and carrion crows.

(2a) in this situation the presence of an empty egg shells near the nest would endanger the brood.

(2b) as this "betrayal effect" decreases rapidly with an increasing distance between eggs and shells this problem is solved by carrying the egg shell away.

----

(3) that's why$_2$ black-headed gulls remove the empty egg shell rather than leave it near the nest.

8.2.5 Appeal to physical laws

There is still one element missing in my account of design explanations (functional explanations$_2$), namely appeal to physical laws. Many design explanations, especially in functional morphology, derive requirements from the laws of physics and chemistry. A typical example is Krogh's derivation of why "larger" animals need a system of convection in addition to diffusion from Fick's law of diffusion. This explanation can be schematically represented as follows:

(1) According to Fick's law of diffusion the rate of diffusion depends on the distance of diffusion, the area available for diffusion, the concentration gradient of the diffusing gas and the diffusion coefficient of the medium in which diffusion takes place according to the following formulae: $J = -D\,A\,\mathrm{d}P/\mathrm{d}x$.

(2) The need for oxygen in vertebrates is such and such, the diffusion coefficient for animal tissue is $2*10^{-5}$ mm^2/atm*s, etc.

(3a) If the distance between the inner organs and the periphery is more than 1 mm vertebrates would not be able to meet their demand for oxygen if they had to rely on diffusion alone.

(3b) the presence of a system of convection would solve this problem.

(4) in Vertebrates the circulatory system provides a system of convection in addition to diffusion.

----

(5) that's why Vertebrates have a circulatory system.

This explanation starts with a physical law (1), next (2) it interprets the quantities in this law in terms of the organism (the oxygen supply at a certain point in the body of an organism is determined by the rate of diffusion, for an organism that has to rely on diffusion alone the relevant distance is that between the organs and the periphery and so on) and fills them out. From this it is calculated that an organism that lacks the trait in question would have a certain problem (3a), and that the presence of the trait in question solves this problem (3b/4).

8.3 Fundamental relations

8.3.1 Introduction

In the previous section I discussed the structure of design explanations (functional explanations$_2$). In this section I discuss the kind of relation with which such explanations are concerned. My two main claims are:

(1) design explanations are concerned with relations of functional interdependence between the parts, processes, behaviours of individual organisms and the environment in which they live (not with the causes of traits);

(2) relations of functional interdependence are constitutive of the structure of the living world.

As said before, my main question is the question what design explanations add to our knowledge. In order to answer this question it is worth paying some attention to the question what causal explanations add to our knowledge. Imagine a situation in which you know all events that happened in a certain space, in a certain period of time. Among other things you know that last week Roger woke up at seven except on Sunday, when he woke up at nine. What insight would be gained if someone causally explained the fact that on Sunday Roger woke up at nine rather than at seven by pointing out that his alarm clock broke down? The insight that there is a connection between the buzzing of the alarm clock and Roger's waking up, of course. You learn that the buzzing of the alarm clock is one of the factors that brings about Roger's waking up. More generally, it seems that causal explanations show us how a certain event, state or capacity hangs together with another part (event, state or capacity) of the world.

This is the position defended by Wesley Salmon in, among others, his *Scientific Explanation and the Causal Structure of the World* (1984). On Salmon's view explanations show us how a certain event fits into the causal structure of the world. As he puts it later:

> [a causal explanation] seeks to provide a systematic understanding of empirical phenomena by showing *how* they fit into a *causal* nexus (Salmon 1989: 120).

Salmon's causal theory of explanation is a special version of a more general view which is called "the ontic conception of explanation". According to the ontic conception explanations show us how a certain phenomenon fits into the structure of the world (causal or otherwise):

> According to the ontic conception [...] the explanation of events consists in fitting them into the patterns that exists in the objective world [...]. We all maintain that explanations reveal the mechanisms, causal or otherwise, that produce the facts we are trying to explain (Salmon 1989: 121).

Proponents of the causal version of the ontic view of explanation (such as Salmon) maintain that the structure of the objective world consists of causal relations only and that, hence, explanations must be causal; proponents of other versions of the ontic conception (such as Coffa and Railton) are more lenient towards non-causal explanations.

The view I will defend is kindred to the ontic conception. I argue that design explanations contribute to our knowledge because they show us how a certain trait of an organism relates to the other traits of that organism and the state of the environment in which it lives. However, the relation is not in terms of causes but in terms of what is useful for the organism to have.

On my view, explanations, in general, aim to show us how a certain phenomenon fits into the structure of the world. That is, they show us how that phenomenon hangs together with the rest of the world (as we know it). Of course, not any relation between phenomena is part of the structure of the world. For example, the relation '... is longer than ...' is not part of the structure of the world: if one shows that Peter is longer than Mary one does not show how Peter and Mary hang together. The question what kind of relations constitute the structure of the world is a scientific issue, that is it is answered in the course of scientific enquiry. The best indication that a certain kind of relation is part of the structure of the world is that it enables us to reach our practical and cognitive purposes. For example, we learn that causal relations are constitutive of the structure of the world both because they are the kind of relations which enable us to influence the course of events and because they enable us to build a unified science.

Proponents of a causal theory of explanation (e.g. Salmon 1984) maintain that the structure of the world is of a causal nature. I argue that, to our best knowledge, relations of functional interdependence too are part of the structure of, at least, the living world.[9] My main argument is that these relations determine which organisms are physical possible (i.e. able to survive and reproduce) and which not. For example, organisms in which the distance between the inner organs and the periphery is greater than 1 mm are not viable if oxygen transport is by diffusion alone. Such organisms must have an active transport mechanism: if they lack such a system they will die. In other words, the possession of a circulatory system makes it possible that an

---

[9] Functional interdependencies are part of the structure of the world of artefacts, too. However, the investigation of artefacts is beyond my subject.

organisms becomes larger. We learn that such relations are constitutive of the structure of the world when we try to build an artefact. Whereas causal relations are the kind of relations that determine what things can be brought about, functional interdependencies are the kind of relations that determine what constructs are stable. Functional interdependencies are the kind of relations that enable us to build stable constructs.

In section 8.3.2 I describe the received view of the structure of the living world. This view distinguishes two kinds of causes: proximate causes which explain how organisms develop and maintain themselves and ultimate causes which explain why and how the different organisms evolved. The *locus classicus* for this view is Ernst Mayr’s “Cause and Effect in Biology” (1961). In section 8.3.3 I propose to replace this dichotomy by a tripartition consisting of (i) causes at the individual level, (ii) causes at the population level and (iii) functional interdependencies. In section 8.3.4 I attempt an account of the relation of functional interdependence and in section 3.5 I discuss the relation between these three kinds of relations.

8.3.2 Mayr’s distinction between proximate and ultimate causes.

Many evolutionary biologists tend to divide biology into two kinds: two fields or modes (“functional biology” and “evolutionary biology”), dealing with two different types of questions (“how questions” and “why questions”) and two different types of causes (“proximate causes” and “ultimate causes”). For example, Futuyma (1986: 286), the “bible” of evolutionary biology, says:

> The study of biology may be divided into two modes. ‘Functional biology’ asks how an organism works—how does it develop and maintain itself? The other approach to biology, the historical approach, asks ‘how and why has life come to be this way?’

The *locus classicus* for this division is Ernst Mayr’s “Cause and Effect in Biology” (1961). In this influential paper Mayr contends that “the word *biology* is a label for two largely separate fields which differ greatly in method, *Fragestellung*, and basic concepts” (p. 1501). Functional biologists are said to ask ‘how’ questions and deal with proximate causes; they study physiological processes, favour the reductionist approach and reach their conclusions by means of experimentation. Evolutionary biologists are said to ask ‘why’ questions and deal with ultimate causes; they study biological diversity and use the comparative method.

This view of biology has its historical roots in the successful attempt of Dobzhansky, Mayr, Simpson and Stebbins to frame the study of evolution as a professional science. Among other problems they had to deal with the prevailing view that biology should aim to reduce biological phenomena to physical processes, if it was to be taken seriously as a science. The proximate / ultimate distinction was Mayr’s answer to this problem. This distinction should at once confirm the autonomous character of biology and the legitimacy of evolutionary biology as a science.

Science according to Mayr aims to discover the causes of what happens. However, in the living world almost any phenomenon has two sets of causes: a proximate set and an ultimate set. The proximate set is the object of traditional biology (physiology or functional biology, as Mayr called it), the ultimate set was to be the object of the new study of evolution. A complete explanation would identify both sets, and both studies are therefore equally legitimate.

As Mayr repeatedly emphasizes neither the terms 'proximate cause' and 'ultimate cause' nor the idea that biology has two sides were new. Mayr, however, was the first who attempted a more systematic treatment of these concepts and who made them into the foundation of biology.

Mayr introduces his proximate / ultimate distinction by means of an example, namely the question why a certain warbler individual started to migrate southward on August 25th (Mayr 1961: 1502/3). Mayr lists four causes for this migration:

(1) "an intrinsic physiological cause": in response to a decrease in day length the warbler became physiologically ready to migrate in response to certain weather conditions;
(2) "an extrinsic physiological cause": the weather conditions on August 25 were such that the warbler, already physiologically ready for migration, started off that day;
(3) "a genetic cause": in the course of evolution the warbler has acquired a genetic constitution that induces the appropriate physiological response to environmental stimuli;
(4) "an ecological cause": "the warbler, being an insect eater, must migrate, because it would starve to death if it should try to winter in New Hampshire".

According to Mayr "we can readily see" that these causes fall into two types. The physiological causes might be called the proximate causes; the other two the ultimate causes.

Note that Mayr's use of the term 'cause' is ambiguous. The term is often more or less synonymous with 'explanation' (or with 'causal explanation') but in some places it is better interpreted as meaning 'the conditions that brought about a certain effect' (that is as 'cause' in a more usual sense).

In order to clarify the distinction between proximate and ultimate causes Mayr invokes the notion of a 'genetic program'. According to Mayr proximate causes are the explanations that deal with the decoding of the genetic program. Ultimate causes, on the other hand, are "the causes that have a history and that have been incorporated into the system through many thousand generations of natural selection". I assume that he means that ultimate explanations deal with the history of the genetic program.

The notion of a 'genetic program' is controversial. Mayr does not explain this notion and answers the objection that this notion is unclear by saying that it is well established in biology. This answer is unsatisfactory but that need not distract us. Mayr's distinction can be made without using the notion of a genetic program: proximate explanations are concerned with the mechanisms that bring about a certain reaction in a certain individual; ultimate explanations seek to explain the differences between the mechanisms of individuals of different taxa. In later work

Mayr uses the terms “ultimate”, “evolutionary” and “historical” as synonyms (e.g. Mayr 1982: vii, 67) as do his followers. Brandon (1981: 93) for instance, states that ultimate explanations answer the question “how and why has this mechanism evolved”. In a recent book Mayr says that “evolutionary (historical or ultimate) causes [...] attempt to explain why an organism is the way it is, as a product of evolution” (Mayr 1997: 117).

Table 3.1: design explanations do not fit into Mayr’s classification of explanations

proximate explanation (Mayr)	ultimate explanation (Mayr)	design explanation (as I have shown)
concerned with proximate causes	concerned with ultimate causes	concerned with utilities
functional biology	evolutionary biology	functional biology
physiological processes	biological diversity	(differences in) form, activity, behaviour
answers how-question	answers why-question	answers why-question
reductionist	historical	holistic
employs physical sciences	employs population genetics	employs physical sciences
concerned with the decoding of a genetic program	concerned with the history of a genetic program	concerned with the interrelation between the parts and behaviours of organisms and their environment
how things work	how and why things evolved why things are the way they are	how things hang together why things are the way they are
experimentation	comparison	comparison, experimentation and calculation

Mayr's insight that there are two equally legitimate and complementary ways to fit biological phenomena into the causal structure of the world has been a major breakthrough in the philosophy of biology. In my proposal in the next section I take this insight into account as the distinction between individual level causal explanations and population level causal explanations. However, Mayr’s account fails to do justice to design explanations (see table 3.1) and confuses them with evolutionary explanation.

As I have argued design explanations address the questions why an organism is as it is, but they answer this question not in terms of their evolutionary history but in terms of what is useful to survive and reproduce. Design explanations are also used to explain the differences between the mechanisms in individuals of different taxa, but they point to the different needs which the different mechanisms should satisfy, not to the different histories of those mechanisms. Design explanations are holistic in character (they explain individual level traits in terms of other individual level traits) but they are established by means of experimentation and calculation, and they show an orientation towards the physical sciences.

Mayr's own example of an ecological cause. According to Mayr ultimate causes are “the causes that have a history and that have been incorporated into the system through many thousand generations of natural selection”. However, Mayr's example of an ecological cause is a clear case of a functional counterfactual: “the warbler, being an insect eater, *must* migrate, because it *would* starve to death *if* it should try to winter in New Hampshire”. It relates the warbler's migration (the trait to be explained) to its eating habits (“being an insect eater” is the explaining trait) in terms of needs (“the warbler would starve to death”). The truth of this counterfactual is established by experimentation and calculation and has nothing to do with how the warbler and its traits evolved historically. The view of cold winters as the ultimate and historical cause of migration easily gives rise to the misunderstanding that warblers evolved the habit of migration as the result of thousands of years of insect eating in a seasonal climate. Of course, this scenario is wrong. The warbler's ancestors must have been migrating either before they started to live exclusively on insects or else before they moved to a seasonal climate. The fact that an insect eater would starve to death if it did not migrate, however, only explains why insect eaters must migrate (if they are not to starve) not why migrating insect eaters evolved. It seems therefore appropriate to distinguish design explanations from evolutionary explanations rather than to lump them together under the banner of “ultimate causation”.

8.3.3 My proposal

In order to accommodate design explanations I propose to classify the relations to which explanations in biology appeal (the relations that constitute the structure of the living world) along two dimensions: (1) individual level / population level relations, and (2) causes / interdependencies. This results in the following tripartition: (i) causes at the individual level, (ii) causes at the population level, and (iii) functional interdependencies at the individual level. The question whether there exist functional interdependencies at other levels than the individual (e.g. at the population level or at the level of the ecosystem) and what their role is in biology is irrelevant to my subject and for that reason ignored.

Note that, in contrast to Mayr, I clearly distinguish between explanations and the relations with which those explanations are concerned.

Individual / population

The importance of the distinction between relations at the individual level and relations at the population level has been stressed recently by Ernst Mayr, Richard Lewontin and Elliott Sober. These authors emphasize that the insight that evolution is essentially a population phenomenon is one of the major breakthroughs Darwin brought about. Lewontin (1981, 1983) points out that Darwin conceived of evolution in a manner that is radically different from conceptions of historical change before Darwin. Lewontin distinguishes two ways of conceiving of historical change. Before Darwin theories of historical change were *transformational* in nature. That is, they accounted for evolution in terms of individual transformations. For instance, in Lamarck's theory of evolution species change because the organisms in them are gradually modified. Darwin's theory on the other hand is *variational* in character. In this view it is the population rather than the individual that evolves. The population is made up of individuals that differ from each other in some properties and the population evolves by a change in the proportions of the different variants (whereas the individuals may remain the same). (See also Sober 1984: 147-155).

Darwin's emphasis on populations and on individual differences was first brought to our attention by Mayr in his essay "Typological versus Population Thinking" (1959). Mayr summarizes Darwin's innovation as follows:

> For the typologist, the type (*eidos*) is real and the variation an illusion, while for the populationist the type (average) is an abstraction and only the variation is real (Mayr 1976: 28).

Sober (1984: 155-169) elaborates on Mayr's ideas. Following Mayr, he distinguishes two views on variation. In the "typological" or "essentialist" view variation between organisms is conceived of as the result of forces acting on the individual. These forces drive the individual away from a natural state, thought to be typical of a species (the natural state is called the "type" of that organism or of that species). In the Darwinian view, on the other hand, variation is a natural state of populations rather than a deviation of individuals. The discipline that studies the principles of variation is called "population genetics". Population genetics employs a natural state model. This model is the Hardy-Weinberg model, which describes how a population behaves in the absence of "disturbing forces" like selection and drift. This Darwinian view differs from the typological view in several aspects. First, variation itself is viewed as a natural state rather than as a deviation from the natural state. Second, natural states are states of populations rather than of individuals. Third, evolution essentially involves processes (such as selection, drift and hybridization) that work at the population level in addition to processes that work at the individual level (such as mutation and recombination).[10]

Causes / interdependencies

Design explanations differ from causal explanations in that they are not concerned with how a certain event, state, trait or capacity is brought about but with functional interdependencies. Functional interdependencies are synchronic relations which determine which combinations of characters are stable (can survive and reproduce) and which not. It was, once more, Darwin who paved the way to the insight that design explanations do not explain how a trait came into being. In the pre-Darwinian traditions of Cuvier and Von Baer it was supposed that the needs a trait satisfies causally explain the emergence and maintenance of that trait at the individual level. The biologists working in these traditions were impressed by the interdependence of the different parts and processes of an organism. The harmony of the different parts of an organism was understood as the result of the interaction of the needs of that organism at the individual level. According to George Cuvier (the founding father of functional morphology) the two basic principles of zoology are the principle of the conditions of existence and the principle of the correlation of parts. The principle of the conditions of existence states that the different parts and processes depend on each other and support each other.

> Since nothing can exist without the reunion of those conditions which render its existence possible, the component parts of each being must be co-ordinated in such a way as to render possible the whole being, not only in itself, but also with regard to its surrounding relations (Cuvier 1817, vol. 1: 6).

As a result there are certain regularities in the design of organisms: certain organs are always found together and from the form of a certain organ one can infer the presence and the form of certain other organs (this is the principle of the correlation of parts). For example, if the alimentary canal of a certain animal is suitable for digesting flesh (and nothing else) the other organs of the animal must also be adapted to this particular process. The organism must be capable to obtain this kind of food and to process it. The animal will be a fast runner. It will have fangs and carnassial teeth, strong claws and muscles, keen sight, fine smell, and a stream-lined body. Such correlations are to be determined by means of comparative anatomy. According to Cuvier all factual correlations are the result of functional interdependencies. Some correlations, such as that of a carnivorous alimentary canal and strong claws, are explicable on “rational grounds”, which means that by means of plausible reasoning one is able to show that an animal having a

---

[10] In population genetics evolution is often defined as a change in the gene pool of a population over time. The gene pool is the set of all alleles ("genes") in a population. Some authors (e.g. Endler 1986) have objected to the population genetic definition of evolution, but they do not the deny that evolution is a population phenomenon.

certain trait (e.g. a carnivorous alimentary canal) needs certain other traits (e.g. claws). But even correlations that are not explicable on rational grounds, such as the correlation between being ruminate and having cloven hoofs, must be accepted as functional interdependencies on "empirical grounds".

The question 'how did the different parts and processes of an organism became geared to each other?' became one of the central theoretical issues in eighteenth and nineteenth century biology. As is well-known, in British Natural Theology this question was answered by appealing to the hand of a benevolent creator. In main stream biology this doctrine never made headway. Cuvier saw the harmony between the different parts and processes of an organism as the result of the causal interaction of the interdependent parts at the individual level. This interaction was assumed to be a kind of material exchange, called "tourbillon vitale" ("Stoffwechsel" in German). Interdependent parts were thought to maintain each other by means of this "tourbillon vitale". The same process operates in development and regeneration. Cuvier and his followers tend to confuse this assumed causal interaction between functionally interdependent parts with the relation of functional interdependency itself. In their view the fact that several organs are functionally interdependent maintains the gearing between those organs. The needs of an organism (such as the need for strong claws in an organism capable of digesting only flesh) act as efficient causes which organize the process of material exchange in such a manner that the organism's needs are satisfied. Hence, design explanations (which appeal to the need for a certain structure) are seen as causal explanations, they explain how the harmony between the parts and the processes of an individual organism is maintained.

The process of material exchange was thought to explain (in principle) how the harmony of an organism is maintained. The origin of this harmony is another issue. According to these biologists causal interaction in the organic world differs from causal interaction in the non-organic world. Causal interactions in the non-organic world were supposed to be linear, (A->B->C->D), causal interactions in the organic world are "clearly" cyclic (A->B->C->A). In the view of late eighteenth, begin nineteenth century science it is impossible to explain how such a cyclic arrangement of causes came into being.[11] What we can try, however, is to explain how this arrangement is maintained (in the individual) and modified (in the course of the ontogeny) given the fact that there is such an arrangement.

This is were Darwin comes in, or rather his theory. Darwin, himself had a poor grasp of the theoretical issues in contemporary morphology, among others because he had no formal training in biology. He also had difficulties in understanding German, which was the language of biology in that time. Darwin addresses the issues raised in the tradition of natural theology and focuses on biogeographical, taxonomic and ecological questions, rather than on theoretical

---

[11] This view is most clearly expressed in Kant's *Kritik der Urteilskraft* (1790).

morphology. When he speaks of ‘adaptation’ he is either concerned with the relation between an organism and its external environment or with the efficiency with which a specialized organ performs its causal role. Darwin seems unaware of the issue of functional interdependencies. At least, he does not explicitly discuss this issue. However, the upshot of Darwin’s work is that the answer to the question ‘how did the parts and organs of an organism became geared to each other and to the environment in which it lives?’ must be sought in the evolutionary history of the lineage rather than in immediate causal interaction between the parts that are in harmony.

In modern biology the metabolic interaction (if any) between two functionally interdependent organs does not explain the gearing of those organs. Harmony between parts and processes is “pre-stabilized” in the genes. For instance, the lungs of birds have a very complicated structure which is needed to enable flight. In the view of Cuvier, Von Baer and their followers, this harmony is established and maintained by a metabolic process operating between the lungs and the wings of the individual that has both items. In the view of modern biology there is no such exchange. In the course of the ontogeny the lungs and the wings acquire their structure independently. The fact that these structures are in harmony is explained by the fact that in the course of evolution the structures of wings and lungs became tuned to each other (due to selection).

For our purposes, the point is that in Darwin’s trail it became clear that it is important to distinguish explicitly not only between individual level and population level relations but also between individual level causal relations and individual level functional interdependencies (needs, demands and utilities).

Classification of explanations

If the two dimensions (individual / population, causal / functional) are combined one gets four kinds of fundamental relations: (1) causal relations at the level of the individual, (2) causal relations at the level of the populations, (3) functional interdependencies at the individual level, (4) functional interdependencies at the population level. Relations of the fourth kind are not relevant to functional biology (which is primarily concerned to explain the form, activity and behaviour of individual organisms) and for that reason, I will ignore this kind of relations. The resulting tripartition of fundamental relations gives one a better grip on the nature of the relations which are the subject of the different kinds of explanation I discussed in chapter 2 (better than Mayr’s bipartition).

Physiological explanations (the kind of explanations which biologists call ‘causal explanations’) are concerned with individual level causal relations. These come in two kinds: explanations that specify causes for a certain type of change, and explanations that specify properties (among which are capacity explanations, that is functional explanations in Cummins’s sense).

Design explanations (the kind of explanations which biologists call ‘functional explanation’) are concerned with functional interdependencies at the individual level. They explain why it is useful to a certain organism that a certain item or behaviour has a certain character or why it is useful to that organism that a certain causal role is performed by relating that utility to the state of that organism and the environment in which it lives. They explain why a trait is useful to certain kinds of individuals, but they do not explain how that trait was brought about in those individuals. Neither do they explain how that trait was brought about in evolution.

Developmental explanations are concerned with causes at the individual level (just as physiological explanations). They show us how the ovum changes into the adult individual and what the mechanisms are that bring about this change.

Evolutionary explanations are concerned with causes at the population level. They explain how the current state of the population was brought about by a series of changes in an ancestral population. They might also explain the mechanism of change.

Table 3.2 shows the differences between the three kinds of explanations which philosopher’s call ‘functional explanations’.

Table 3.2: Different types of so-called “functional explanations” and their characteristics.

	level	**nature**	**strategy**
Selection explanation	population (lineage)	causal relations	historical
Design explanation	individual	functional interdependencies	holistic
Capacity explanation	individual	causal relations	reductionist

8.3.4 Functional interdependencies

The concept of need

Design explanations are virtually ignored in contemporary philosophy of science, and to the extent that they receive attention they have been confused with capacity explanations or with evolutionary explanations. The nature of the relation of functional interdependence has received no attention at all in contemporary philosophy of science. This in contrast to the nature of causation, which has been the subject of heated debate. In this section I offer a first attempt to analyze the nature of functional interdependence. I develop this analysis by means of an example.

My example concerns the shark’s spiracle. The spiracle is an oval opening which appears to lie just before the gill slits. Among biologists it is well known that the spiracle actually *is* a (modified) gill slit.[12] In contrast with the other gill slits, the spiracular slit is used for the *inflow* of water. The answer to the question why$_2$ sharks need a spiracle appeals both to the shark's need for oxygen and to the feeding habits of sharks. To satisfy the shark's need for oxygen a continuous flow of fresh water over the gills is required. Normally this flow is supplied via the mouth, but when the shark has caught a large prey this channel is blocked. The spiracle affords a channel for the inflow of water when the mouth is blocked. As I said, this design explanation explains the presence of a spiracle in terms of the need for oxygen and the feeding habits. It tells us that the spiracle is needed because the size of the prey makes it impossible to supply enough fresh water via the mouth. The traits to which a design explanation appeals are often said "to demand for" the trait in question. For example, the eating habits and the need for oxygen together demand for a spiracle.

However, a spiracle is not the only possible way to meet the demand for a continuous flow of fresh water over the gills. Lampreys manifest a different solution for the same problem. Lampreys attach themselves to their prey as a result of which they are unable to take in water via the mouth. Their solution is to take in water via the nose. Hence, the spiracle and the nose are functional equivalents.

In analyzing the notions 'needed' and 'demands for' I will apply some ideas put forward by John Mackie (1974). Mackie is concerned with the analysis of the notion of cause. Mackie views a causal relation as a necessary succession of events. His central problem is the distinction between necessary (i.e. causal) and accidental successions. Mackie envisages a chestnut which is treated with a hammer. In the example of a causal succession the chestnut lies on a flat stone and the chestnut is observably flatter after it is hit by the hammer than it was before. In the example of an accidental succession the chestnut lies on a red-hot iron plate and explodes at the moment it is hit by the hammer. Consider the question what makes the first sequence necessary and the second one accidental. Mackie says that "every intelligent and unbiased person" will answer this question by referring to the fact that in the first case the chestnut would not have flattened if it were not hit by the hammer, whereas in the second case the chestnut would have exploded anyhow. More generally his conclusion is that we do not speak of a necessary succession if we have no reasons for a counterfactual statement. In the case of causal relations the counterfactual states that in the given circumstances the effect would not have occurred if the cause were absent. In other words, according to Mackie in the circumstances a cause is necessary for its effect.

Mackie emphasizes that his analyses of a cause as necessary in the circumstances does not imply that the cause is indispensable in the circumstances (see also Mackie 1965). The events which are commonly labelled 'causes' are often neither indispensable nor sufficient for their

---

[12] More precisely, it is the branchial opening lying between the mandibular and the hyoid gill arch.

effects. To take this insight into account he defines the notion ‘at least an INUS-condition’. An INUS-condition is an *i*nsufficient but *n*on-redundant part of a condition which is itself *unnecessary* but *s*ufficient for a certain effect”. The central notion of his analysis, however, is ‘at least an INUS-condition’, which is a non-redundant part of a sufficient condition (this notion differs from that of an INUS-condition in that it includes the cases in which the non-redundant part is sufficient, and in which the condition of which the non-redundant part is a part is indispensable). More formally this may be put as follows:

> An event of type A is *at least an INUS-condition* for an event of type *P* if and only if there are certain (usually unknown) conditions X, such that:

(1) AX[13] is sufficient for *P,* and
(2) X alone is not sufficient for *P*

The statement ‘A is a cause of *P*’ implies according to Mackie:

(1) A is at least an INUS-condition of *P,*
(2) both A and X were realized in the circumstances in which *P* occurred,
(3) all possible alternative sufficient conditions for *P* ($Y1$, $Y2$ and so on) were not realized in the circumstances in which *P* occurred.

A is non-redundant because X alone is not sufficient for *P* but not necessary indispensable because $Y1{=}KX$ might be another sufficient condition for *P*.

In the Hempel and Nagel approach, the question whether a trait is necessary or not is equivalent to the question whether the trait is indispensable or not. I think it is reasonable to say that needed traits are necessary although they are not indispensable. Needed traits are necessary in the sense that if the organism *as it is* would lack the trait it would not be able to survive and reproduce. More formally this may be put as follows:

> A trait A is *needed* in an organism *i* if and only if in *i* or in *i*’s environment a certain condition X is realized, such that:

(1) the realization of both A and X is sufficient for *i* to survive and reproduce, and
(2) X alone is not sufficient for *i* to survive and reproduce, and
(3) for all Y A:[14] the realization of Y and X is not sufficient for *i* to survive and reproduce[15], and
(4) possible alternatives for A ($K1$, $K2$, etc.) are not realized in *i.*

---

[13] AX means: both an event of type A and condition X are realized. This condition is satisfied if an event of type A occurred.

[14] ‘Y A’ means that A is a combination (“conjunction”) of conditions and Y is one of the parts (“conjuncts”).

[15] This condition is added in order to avoid that the combination of a needed condition (e.g. the presence of a spiracle in sharks) and an unneeded one (e.g. the white colour of its bones) counts as needed.

A is necessary because X alone is not sufficient for *i* to function adequately (a shark would die if it lacked a spiracle), but possibly dispensable because KX might be another sufficient condition (for example, K might be the intake of water via the nose).

This definition explains why counterfactuals have such an important role in design explanations. For example, compare a needed trait such as the spiracle of the shark, with a non-needed one such as the colour of its bones. How do we know that the spiracle is needed and the colour of its bones not? Consider an experimental set up in which the shark is fed large preys only. In this situation, if one would block the spiracle the shark would die, but if one would paint its bones purple this would not have any consequence for its survival. Hence, the need for a spiracle is revealed by an argument and/or an experiment which shows that the shark could not survive if it lacked the spiracle.

The relation 'demands for'

The relation 'B demands for A' is more difficult to characterize. The basic idea is that B demands for A if B is at least an INUS condition for the utility of A. Compare the relation between the feeding habits of the shark and the presence of a spiracle, with an accidental "relation", such as that between the cartilaginous character of the skeleton and the presence of a spiracle. The eating habits of the shark demand for a spiracle, but the composition of the skeleton is irrelevant to the need for a spiracle. Again the difference is expressed in terms of a functional counterfactual: if the shark did not eat large preys, a spiracle would not be needed, but the composition of the skeleton is irrelevant to the need for a spiracle.

Note, that in contrast with causal relations, such as 'brings about', the relation 'demands for' is not transitive. For example, the lack of small preys demands for a shark's habit to catch large preys and the habit to catch large preys demands for a spiracle. However, the lack of small preys does not (directly) demand for a spiracle: if there were enough small prey available and the shark would catch large prey, a spiracle would still be needed. A related difference between causal relations and functional interdependencies is this. Causes form a chain and the same event never occurs twice in a chain of causes (there is no backward causation), but functional interdependencies form a network and although most connections will have a direction (the habit to catch large preys demands for a spiracle but not the other way round) there is no direction at the global level. In order words: the transitive closure of the causal relation is a partial ordering but the transitive closure of the relation 'demands for' not.

As a first impulse one might want to spell out the idea that a certain trait demands for another trait in the following way:

First attempt. A certain condition B (realized in an organism i or in the environment in which i lives) *demands for* trait A of an organism i[16] if and only if in i or in i's environment a condition X is realized, such that:

(1) the realization of A, B, and X together is sufficient for i to survive and reproduce,
(2) B and X *together* are not sufficient for i to survive and reproduce,
(3) for all $Y \quad A$: the realization of Y, B, and X is not sufficient for i to survive and reproduce, and
(4) the realization of X alone is sufficient for i to survive and reproduce.

According to this definition the habit to eat large preys demands for a spiracle, but the cartilaginous character of the skeleton does not demand for a spiracle.[17] Which is desired.

However, this definition does not work for several reasons. One is that B itself might be needed in which case X alone might not be sufficient for i to survive and reproduce. In other words, if someone says that in i B demands for A this does not imply that i could live without A if B were not present. This problem is solved by the following definition:

Second attempt. A certain condition B (realized in an organism i or in the environment in which i lives) *demands for* a trait A of an organism i if and only if in i or in i's environment a condition X is realized, such that:

(1) the realization of A, B, and X together is sufficient for i to survive and reproduce,
(2) B and X *together* are not sufficient for i to survive and reproduce,
(3) for all $Y \quad A$: the realization of Y, B, and X is not sufficient for i to survive and reproduce, and
(4) for all Z $(A\ /\ Z, B\ /\ Z)$: if an organism in which A, X and Z are realized is able to survive and reproduce then an organism in which X and Z alone are realized is able to survive and reproduce.

This definition says that "B demands for A" implies that A would be redundant if B were absent.

However, if there is a Z $(A\ /\ Z, B\ /\ Z)$ such that AXZ is able to survive and reproduce and a part of Z demands for A then condition (4) is not satisfied. This is unwanted. For example, in the case of the shark the habit of eating large prey demands for a spiracle. Yet, there is a Z such that AXZ is able to survive and reproduce and XZ not: a bottom-dwelling shark without a spiracle would have problems maintaining the flow of water over the gills if it lays down on the

---

[16] This is meant to imply that A is present in i.

[17] Suppose A is the presence of a spiracle and B the cartilinagous character of the skeleton. If condition (2) applies BX is not able to survive if the spiracle is absent. Since, the shark would be able to survive if it did not eat large prey, this means that X must include a condition that makes a spiracle necessary, such as the habit of eating large prey. But in that case (4) does not apply.

sand (the mouth lays on the sand then). This means that a bottom-dwelling shark with a spiracle is able to survive and reproduce, but a bottom-dwelling shark without a spiracle not. This example is not far-fetched. Skates and rays spend a large part of their life on the bottom of the sea, with their mouth more or less buried in sand or mud. These creatures have a very large spiracle which meets the need to maintain the inflow of clear water when they are buried. It appears that condition (4) is too strong. Apparently, 'B demands for A' means that B is part of a disjunction of conditions each of which might create the need for A. This brings me to the following attempt:

Third attempt. A certain condition B (realized in an organism *i* or in the environment in which *i* lives) *demands for* a trait A of an organism *i* if and only if there is a set of conditions U ($B \quad U$)[18] and in *i* or in *i*'s environment a condition X is realized such that

(1) the realization of A, any U_i ($U_i \quad U$), and X together is sufficient to survive and reproduce,

(2) the realization of any U_i ($U_i \quad U$) and X together is not sufficient to survive and reproduce,

(3) for all Y ($Y \quad A$): the realization of Y, B, and X is not sufficient for *i* to survive and reproduce, and

(4) for all Z ($A \ / \ Z$, no $U_i \ / \ Z$): if an organism in which A, X and Z are realized is able to survive and reproduce then an organism in which X and Z alone are realized is able to survive and reproduce.

This definition too has a problem. It might be the case that a certain trait is needed for more than one reason and that those reasons are independent of each other. Recall, the example of inquilinism in pearlfishes (example 4.1 of section 2.2.4). This behaviour is needed for two reasons: it is needed to avoid predators and it is needed to avoid light. This means that there are two conditions which independently demand for the inquilinistic behaviour, namely the presence of a certain kind of predators (U_1) and the absence of pigment (U_2). If one or both of those conditions itself is needed (1) does not apply. This brings me to my final attempt:

Fourth attempt. A certain condition B (realized in an organism *i* or in the environment in which *i* lives) *demands for* a trait A of an organism *i* if and only if there is a set of conditions U ($B \quad U$) and in *i* or in *i*'s environment a condition X is realized such that

(1a) the realization of A, all U_i 's realized in *i*, and X together is sufficient to survive and reproduce,

---

[18] '$B \quad U$' means that U is a disjunction of conditions and B is one of the disjuncts (B is a U). B might be the only (element of) U.

(1b) the realization of A, any U_i (U_i U and U_i not realized in A), and X together is sufficient to survive and reproduce,

(2) the realization of any U_i (U_i U) and X together is not sufficient to survive and reproduce,

(3) for all Y (Y A): the realization of Y, B, and X is not sufficient for i to survive and reproduce, and

(4) for all Z (A / Z, no U_i / Z): if an organism in which A, X and Z are realized is able to survive and reproduce then an organism in which X and Z alone are realized is able to survive and reproduce.

There is however one more problem. None of these definitions excludes that a combination of a relevant factor (say the habit too eat large preys) and an irrelevant factor (say the presence of a tail) demands for A (in our example the spiracle). This problem can, I think, be solved by requiring that if B demands for A then any V B demands for A. I am sure that there are further problems once we make the case even more complex. However, it is not the purpose of this chapter (and book) to work out this problem in further detail. The previous general analysis should suffice for the moment.

8.4 Design explanation and functional interdependencies

8.4.1 What is learned from a design explanation

Functional biologists are concerned with three kinds of fundamental relations: causal relations at the individual level, causal relations at the population level and functional interdependencies at the individual level. Design explanations are concerned with the latter kind of relations. Design explanations show their audience how the different parts of an organism, the form and activity of those parts, the behaviour of the organism are functionally dependent on each other and on the environment in which the organisms lives. I will discuss a number of examples to support this claim.

Consider for example the explanations concerned with circulation and respiration (section 4.2.3 and 6.2.2). These explanations show how the presence of a system of blood circulation with respiratory pigments relates to the physical characteristics of the diffusion process, to the size of the organism in question, to its need for oxygen (and therefore to its activity), to the physical characteristics of the blood, to the capacities of the pump and so on. They show how the form of a flatworm relates to the absence of a circulatory system. They relate the presence of respiratory organs and the form which they take to the size of the organisms in question, their need for oxygen, the physical characteristics of the environment in which they live, and so on.

Similarly, Habibi, Thouless and Lindsay (1993) (example 3.3 of chapter 2) relate behavioural differences between sand gazelles and mountain gazelles to each other, to the laws of probability, to physiological and physical characteristics of new-born gazelles and lactating offspring and so on (see fig. 81.)

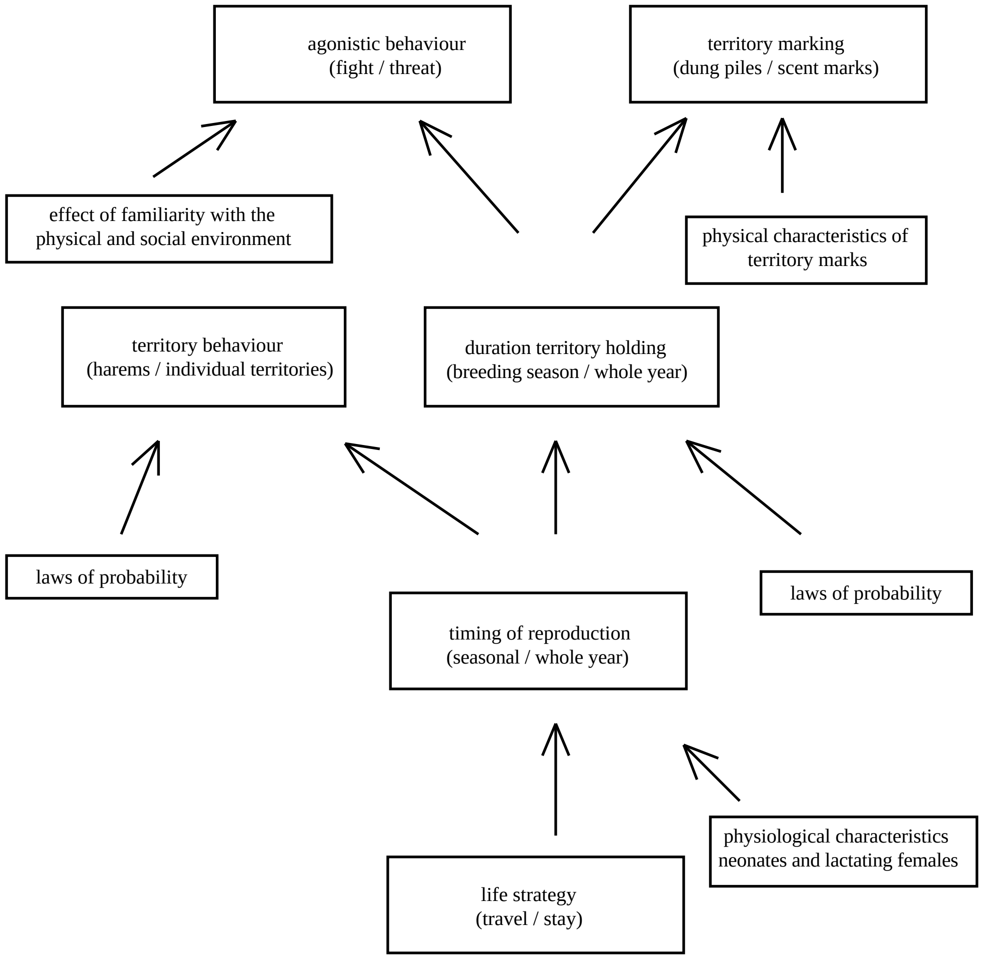

Fig 8.1: Differences between sand gazelles and mountain gazelles

In section 8.2 I discussed the structure of design explanations. Many design explanations are much more complex than the examples which I discussed in that section: typically they combine statements of several types. I now discuss three typical examples of complex design explanations. These examples connect the discussion of the structure of design explanations in section 8.2 to the main point of this section, namely that design explanations show us how things hang together in terms of what is useful to survive and reproduce. Furthermore, they illustrate three strategies to develop design explanations: (1) explain how a certain utility criterion is optimized or how a certain requirement is met, (2) pose a problem that would occur if the trait in question were replaced by an alternative trait and explain why this problem is a problem, (3) identify a complex task in which the item or behaviour the structure of which is to be explained takes part, make up a simple implementation and explain why this implementation does not work in the conditions in which the organisms in question live, pose a solution for this problem and repeat the process until one has developed a implementation similar to the structure to be explained.

Why mammals gallop

A straight forward way to explain why an organism is built the way it is built or why it behaves the way it behaves is this: pose an optimality criterion or a requirement that applies in the conditions in which the organisms in question live and show that an organism with the traits in question scores better on this criterion than an organism with the alternative traits or that an organism with the traits in question meets this requirement whereas an organism with the alternative traits would not meet the requirements. An example of such an explanation is McNeill Alexander’s (1988) explanation of why mammals gallop. Most larger mammals (such as dogs, horses and antelopes) change gaits as they increase speed. At low speed the gait of these mammals is symmetrically: the left and the right leg of a pair swing in opposite directions (such as in walking and trotting). At high speeds they change to galloping, an asymmetrical trait in which both the forelegs swing in the opposite direction as the hindlegs. McNeill Alexander and his colleagues expected that large mammals gallop because galloping minimizes the energy costs of fast locomotion. The first attempts to work out this idea by means of mathematical modelling seemed to refute it: the models appeared to show that, even at high speeds, there always is a symmetrical gait that is at least as economically as galloping. However, metabolic measurements showed that in the range of speeds at which horses normally gallop, galloping is *de facto* more economically than any other gait. It remained a riddle why this is the case, until the researchers realized that the aponeurosis of the backbone could serve as a strain energy store. As an animal increases its speed, the fluctuations of the internal kinetic energy of the legs become larger. At a certain speed it becomes useful for the organism to balance these fluctuations by storing strain energy. McNeill Alexander and his colleagues had previously rejected the possibility that elastic items could help to save energy by balancing energy fluctuations. They now show that their previous objection holds only if the organism’s gait is symmetrical. If the gait is asymmetrical (such as in galloping) the aponeurosis can serve to store energy. Thus, the explanation of why large mammals change from symmetrical gaits to galloping at high speeds is probably that this change enables the animal to store strain energy in the aponeurosis, which saves energy at high speeds and only at high speeds.

The train of thought of this explanation can be represented schematically as follows:

(1) It is useful for an organism to perform a certain activity with as less energy costs as possible.

(2) At high speeds it becomes possible for large mammals to save energy by storing strain energy, whereas at low speeds storing strain energy does not pay much.

(3) The aponeurosis of the vertebral column of large mammals serves as a strain energy store.

(4) The aponeurosis can perform this causal role (function$_2$) if the gait is asymmetrical but not if it is symmetrical.

(5) Galloping is an asymmetrical gait whereas trotting and walking are symmetrical

----

(6) That's why large mammals gallop at high speeds rather than trot or walk

This explanation starts by stating a general criterion for what counts as useful, namely to perform a certain activity with as less energy costs as possible (1). Next, it points out a way to implement this criterion (by storing strain energy). It is also emphasized that this implementation would be useful only in the conditions in question (at high speeds) (2). Then, it points out an item that implements the causal role mentioned in (2), namely the aponeurosis of the vertebral column (3). Sentence (4) states a requirement that applies if the item is to perform that causal role. Statement (5) points out that in the condition stated in (2) the trait in question meets the requirement stated in (4).

This explanation relates the habit to gallop to the way in which larger mammals are built, to the conditions in which this habit takes place, and to the laws of mechanics.

Photoreceptor twist

In the example above, McNeill Alexander and his colleagues had a clear idea of the relevant optimality criterion. However, in many cases it is not clear from the outset what the relevant requirements / optimality criteria are. One way to deal with such a situation is to investigate the problems that would result from replacing the traits in question by the alternative traits: what problems would an organism encounter if it had the alternative traits instead of the traits in question and how do these problems relate to survival and reproduction. A typical example of such an explanation is Wehner & Bernard's (1993) explanation of photoreceptor twist in bees. In 1975 Rüdiger Wehner, Esther Geiger and Gary Bernard reported that the majority of the light-sensitive cells in a bee's eye are twisted along their length, just like a cork screw. Initially, this claim was met with disbelief. It was suspected to be an artefact of the method of preparation. However, in later years the researchers were able to establish the existence of the twist beyond any reasonable doubt. In their 1993 paper Wehner & Bernard address the question why$_2$ these cells are twisted. They show that this twist "is *necessary* for reliable encoding of information about color" (p. 4132, emphasis mine). If the cells were not twisted bees would experience a psychedelic world of ever changing colours, in which it would be difficult to impossible to find food.

Insect eyes are made up of multiple facets. Beneath each facet there is a rod-shaped light receptor, the rhabdom. Each rhabdom is composed of eight photoreceptor cells. The photoreceptor cells contain a light sensitive pigment. Due to differences in the pigment they contain, the photoreceptor cells of bees fall into different types, each of which is most sensitive to a distinct part of the spectrum. That is, cells of different types respond differently to different wavelengths. There are three types with a maximum sensitivity in, respectively, the ultraviolet, blue, and green part of the spectrum. These differences form the basis of colour vision in bees. The light sensitive pigments of all insects are stacked in an ordered array of microvilli at the edge of the light-sensitive cells. In a straight rhabdom all the microvilli of a certain photoreceptor cell are oriented in the same direction. This arrangement acts as a polarization filter. As a result the response of the different light-sensitive cells in a straight rhabdom depends not only on the wavelength but also on the angle of polarization of the incoming light. With respect to microvillar orientation there are two types. One is sensitive to light that is polarized in the dorso-ventral (“vertical”) direction, the other to light that is polarized in the lateral (“vertical”) direction.

In real bees most of the rhabdoms in the compound eye are twisted; straight rhabdoms are limited to the dorsal rim. To explain the twist Wehner and Bernard envisage a hypothetical bee in which all the rhabdoms are straight. The sensitivity to polarization angle would create serious difficulties for such bees. Light reflected from the waxy surfaces of plant is polarized. The direction of polarization of light reflected by a particular leaf depends on the angle between the source of light and the plane of the leaf. The angle of polarization as it is perceived by the bee in its turn depends on the direction of polarization of the light and on the line of sight. As a result, the hypothetical “straight only” bee would perceive a change of “colour” whenever it changes its direction of flight. The authors computed the sensitivity of the different types of straight and twisted photoreceptor cells as a function of wavelength. They did so for horizontally and vertically polarized light. In addition they measured the state of polarization of light reflected from plant surfaces. The measurements and calculations show that the perceived colour of the vegetation would change dramatically with the direction of illumination and the bee’s line of sight. As a result of this it would be nearly impossible for the hypothetical bee to recognize and localize its food resources (flowers):

> For example, when zigzagging over a meadow, with all its differently inclined surfaces of leaves, the bee would experience pointillistic fireworks of false colors that would make in difficult to impossible to detect the real color of the flowers (Wehner & Bernard 1993: 4134).

The authors refer to this problem of their hypothetical bee as the “false-colour problem”. (I think this is a rather unfortunate label: the problem is not so much that the colours are false, whatever that may mean, but rather that the perceived “colour” is not independent of the position of the bee and the direction of the incoming light). In a real bee this problem does not arise as the photoreceptor twist eliminates the sensitivity to polarization angle (as the authors demonstrate by means of calculation). In other words:

> The bee’s solution to the false-color problem is to get rid of the polarization sensitivity of its eyes by twisting its photoreceptors (Wehner & Bernard 1993: 4134).

Hence, the authors explain why most of the rhabdoms in the compound eye of bees are twisted by pointing to four factors, namely (1) the fact that differences in sensitivity of photoreceptor cells play an important causal role in the way bees process visual information (“colour” vision), (2) the fact that bees recognize and localize their food-resources (flowers) by colour, (3) the fact that light-sensitive pigments are stacked in an order array of microvilli, (4) the fact that light reflected by waxy plant surfaces is polarized. They argue that the presence of these factors would create a problem for a hypothetical organism in which the trait in question (photoreceptor twist) were absent and that the trait in question solves this problem (see fig. 8.2).

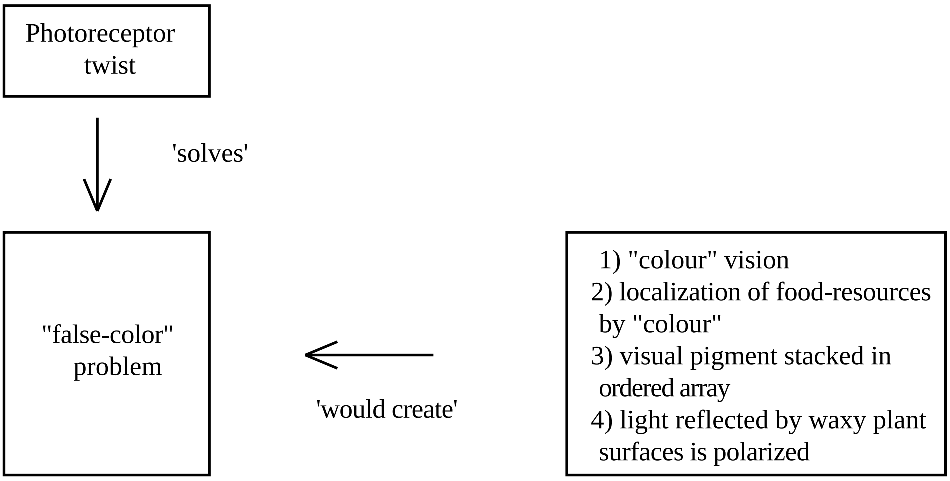

fig 8.2 photoreceptor twist

The problem consists of a series of “unwanted” (p. 4134) effects (fig 8.3). These effects are unwanted because they result in the death of the hypothetical “straight only” bee. In straight rhabdoms, the response of a certain photoreceptor cell depends not only on the wavelength of the incoming light but, due to the manner of pigment stacking (3), also on its polarization angle. Light reflected by the vegetation is polarized (4). This (3+4) means, that the “colours” perceived by a hypothetical bee with only straight rhabdoms would depend on the position of the bee and the direction of the incoming light. This would make it impossible for that bee to recognize and localize flowers by “colour”. Since, bees depend on “colour” to recognize and localize flowers (1/2) the hypothetical bee would starve to death. Twisted photoreceptors are not sensitive to polarization angle and, hence, the problem is not generated in a real bee.

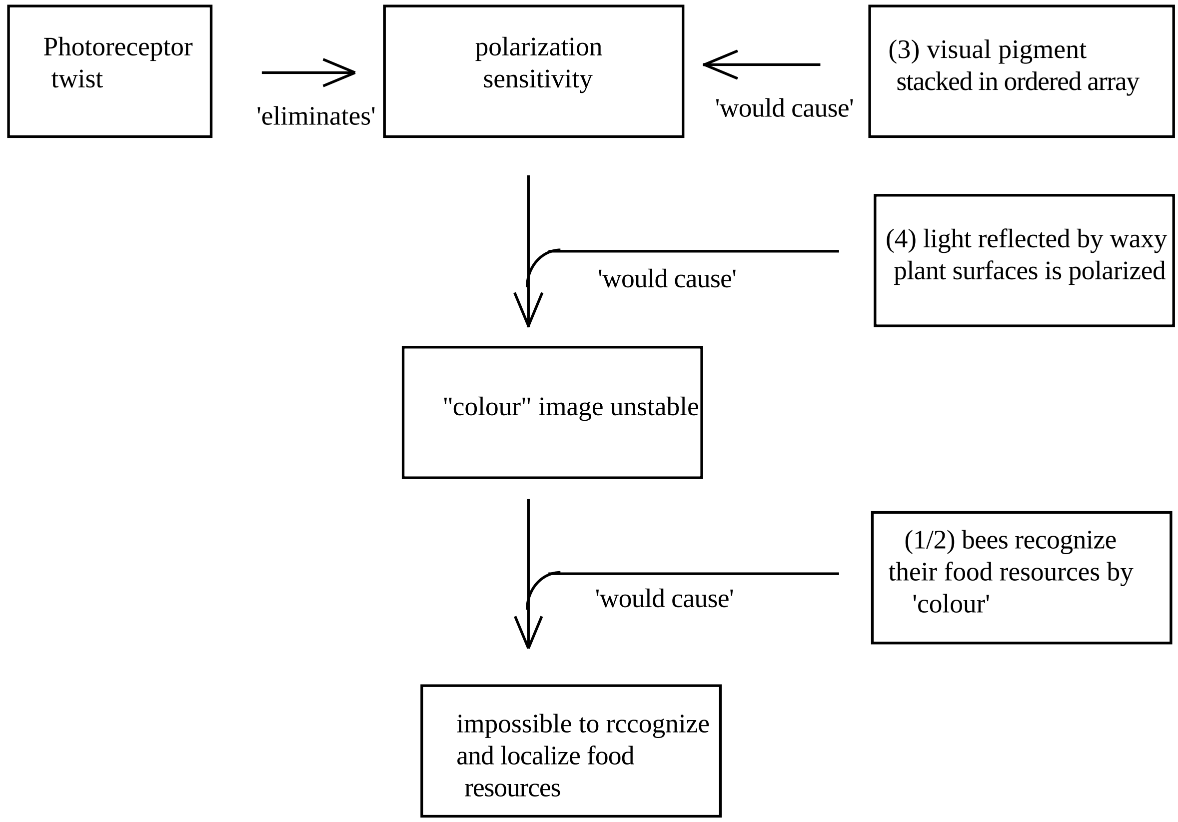

Fig 8.3: The false colour problem and its solution

Note that this explanation is easily converted into one that shows how a certain requirement is implemented (see fig. 8.4). The principle requirement is that given the bee’s feeding habits and the environment of its food resources, the bee’s mechanism of “colour vision” should enable the bee to recognize and localize flowers by sight against a background of green leaves. Given the fact that light reflected by plant leaves is polarized, this is possible if bees distinguish “colour” on the basis of wavelength, but not if “colour” depends on polarization angle. (Note at this point that in addition to explaining the photoreceptor twist the authors explain why bees distinguish “colours” on the basis of wavelength rather than on polarization angle. They do so by showing that wavelength-colours fit the requirement and polarization-colours do not.) Hence, the bee needs to get rid of any determination of “colour” by polarization angle. Given the arrangement of visual pigments a bee would not meet this requirement if all its rhabdoms were straight, but the twist solves this problem, and, hence, satisfies the need to get rid of polarization. The authors summarize their argument in the conclusion that the twist is a "necessary requirement" for recognizing and localizing flowers against a background of green leaves. It will be clear from my account that this conclusion is but a pale shadow of what they have done, namely showing how the twist is situated into a network of needs and requirements created by the other traits of the bee and the environment in which it lives.

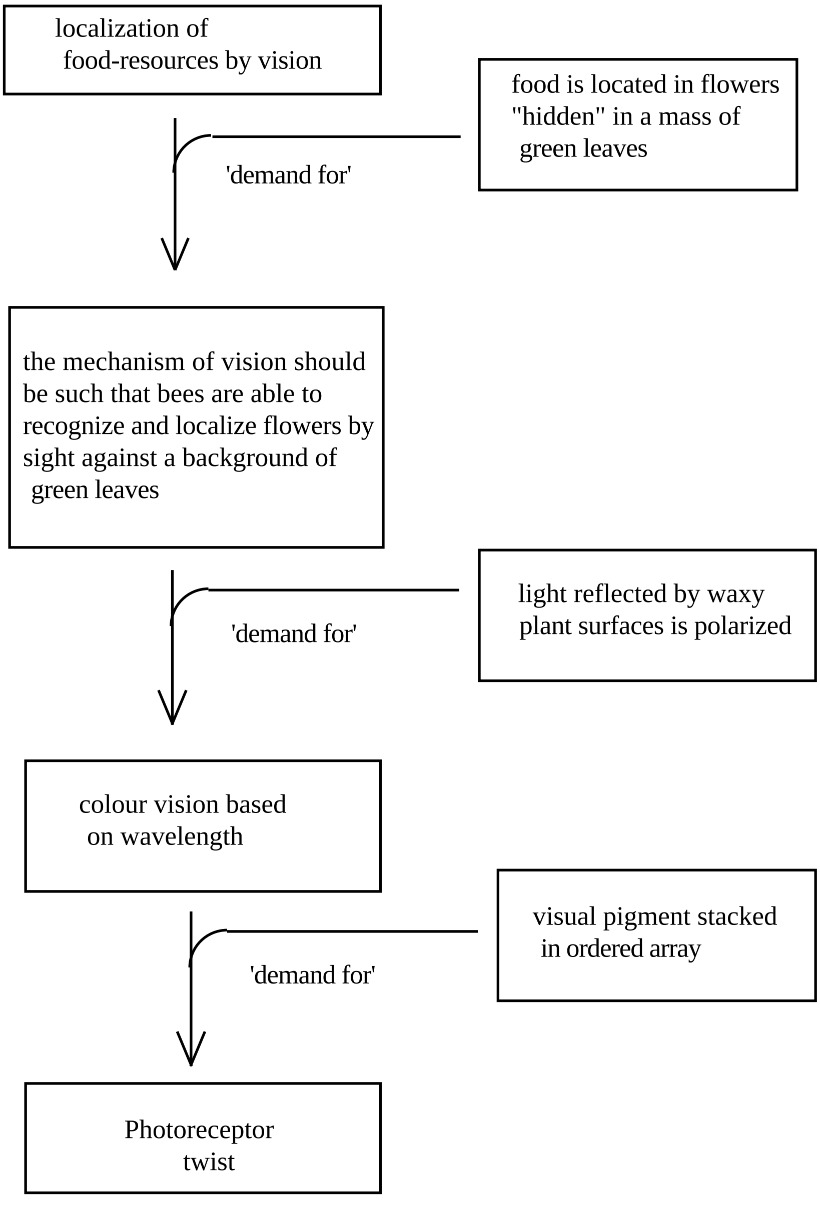

Fig 8.3 Photoreceptor twist

Egg shell removal

In the examples above the explanation focuses on one trait (the tendency of large mammals to change gait at high speeds, respectively the twist in the photoreceptors of bees). If one wants to explain complex structures or behaviours and one has only a vague idea of what the relevant alternate traits, requirements, and conditions are the following strategy is often used:

1) identify a complex task (function$_2$) in which the item or behaviour the structure of which is to be explained takes part
2) describe a simple implementation to perform this task
3) explain why this implementation has problems in the conditions in which the organisms in question live
4) describe a more complex implementation which solves this problem
5) repeat step (3) and (4) until one has an implementation similar to the structure or behaviour in question.

One of the best examples of this strategy is Van der Stelt's (1968) explanation of the arrangement of the muscle-fibres in sharks. Another good example is Langman's (1989) explanation of the vertebrate immune system. These examples are too complex to be discussed here. Instead, I will return to an example I have already discussed and show how this example fits into the third explanatory strategy. This example is the explanation of Tinbergen and his students (1962) of why many birds remove the empty egg shell after the chick has hatched. I will add details I have not mentioned before.

Tinbergen c.s. seek to explain the egg shell removal behaviour of the black headed gull. As many birds, black headed gulls remove the empty egg shell after the chick has hatched, but in contrast to waders such as the ringed plover and the oystercatcher they wait for an hour or two before they before removing the empty shell. Tinbergen c.s. start with the hypothesis that the egg-shell removal behaviour has a causal role in protecting the eggs against predators (step 1). As a first implementation of this task (protecting the eggs against predators) they envisage eggs which are camouflaged by their colour pattern (step 2). Their experiments show that bird predators find chicken eggs as easily when they are white as when they are painted like the eggs of black headed gulls (step 3). Perhaps, it is not only the colour of the eggs but also their size which plays a causal role in camouflage (step 4). To the surprise of the research team their experiments showed that carrion crows and herring gulls find the eggs of black headed gulls that are painted white almost as easily as natural eggs (step 3 repeated). In nature the eggs of black headed gulls are covered with grass and it turned out that natural eggs covered by grass are found less easily than covered eggs which are painted white (step 4 repeated). Next, the team showed that if there is an empty egg shell near the nest, the eggs are found more easily (step 3, repeated again) and that this problem is solved if the empty shells are carried away (step 4 repeated). This explains the removal behaviour but not the two hour delay. To explain the delay the team points out that in the case of colony breeders such as the black headed gulls the removal behaviour has an important disadvantage beside the advantage mentioned above. After hatching, it takes the chick a few hours to dry. In any colony of black-headed gulls there are some individuals who prey selectively on nearly hatched eggs and wet chicks. They take the chick when the parents leave the nest to attack predators or to remove the egg shell (step 3, again). As dry chicks are left alone, this problem is solved by delaying the removal for an hour or two (until the chick is dry) (step 4 again). Waders do not live in colonies and their bills aren't fit for eating chicks. Which explains the lack of a delay in their case.

8.4.2 Relation between design explanations and capacity explanations

Although capacity explanations and design explanations are to be distinguished carefully, they complement each other well. My claim is that one of the basic interests of functional biologists is to explain how the parts and processes of an organism hang together with each other and with the environment in which the organism lives. In exploring that network of interdependencies they use two main entries. One is to ask the type (4b) question "why$_2$ do such and such organisms have a certain element or perform a certain activity?", the other is to ask the type (3) question "how do such and such organisms manage to perform such and such task"? The type (4b) question is answered by means of a design explanation. Such an explanation typically starts with the attribution of a causal role in a complex activity or capacity to the element or activity to be explained. Such an attribution can be used to explain that complex activity or capacity by means of a capacity explanation. The type (3) question is answered by means of a capacity explanation. Capacity explanations attribute causal roles to the parts and subparts involved in that task. Such attributions can be used to explain the character of the parts or subparts in a design explanation. In short, answers to a type (4a) question can often be used in answering a type (3) question and the other way round. For that reason much research in biology deals with two questions at once. A nice example of such a "double-focus" approach is Eliott, Tait and Briscoe's (1993) study of the function of the crural glands of velvet worms. Velvet worms are the only members of one of the minor phyla of the animal kingdom: the onychophora. They are of considerable interest in evolutionary studies because they have characteristics in common with both the annelids and the arthropods. On the one hand the authors address the type (3) question "how conspecifics locate and recognize each other as mates" (p. 1). This question arises because of the habitat and distribution of the velvet worms: "Onychophorans are usually sparsely and patchy distributed in the dark humid crevices within rotten logs, under stones, and in leaf litter soil. Furthermore, sympatry of up to three species has been identified within the same habitat" (p. 1). The authors argue that visual and acoustic communication is improbable. The fact that the surface of the body of velvet worms is well supplied with chemoreceptors suggests that one has to look for communication by means of pheromones. On the other hand, the authors raise the type (4) question why the onychophorans have crural glands (exocrine glands that open on the ventral surface at the base of the legs). Their restriction to males suggest a sexual function. The authors were able to extract a secretion from the crural glands of males of a certain species and they show that this secretion attracts females of the same species. They conclude that their experiments establish “the function of the crural gland [in individuals of that species] as a female sex attractant” (p. 7). They add that “the role of pheromones for communication in onychophorans is clearly of adaptive significance considering their sparse and disjunctive distribution and the cryptic lifestyle characteristic of the group” (p. 8).

8.4.3 Conclusion

There are two kinds of design explanations. One kind is concerned with the type (4a) question why it is useful to individuals of certain species to have a certain item or to perform a certain activity or behaviour. They answer this question by, first, identifying a causal role of the item or activity in question and, next, showing that in the conditions that apply to the organism in question the performance of that role is useful to those organisms. The other kind of explanation addresses the type (4b) question why a certain item or behaviour has the character it has. They answer this question by, first, identifying a causal role of that item or behaviour and, next, showing that in the conditions that apply to the organism in question that causal role is better performed by an item that has the character to be explained than by an item with some other conceivable character. The conditions to which a design explanation appeals are other traits of the organisms in question and the state of the environment in which they live. This means that design explanations are concerned with relations at the individual level, just as physiological and developmental explanations. However, whereas physiological and developmental explanations are concerned with causal relations, design explanations are concerned with functional interdependencies. The main relations of functional interdependency are needs and demands. Roughly spoken, a trait is needed (required) if its presence is at least an INUS condition for an organisms to survive and reproduce. Roughly spoken, something (*B*) demands for something else (*A*) if its (*B*’s) presence is an INUS-condition for the need for that something else (*A*). Design explanations show how the different traits of an organism hang together in terms of how the different items satisfy demands imposed on them by the way in which the organism is built, the way it behaves, the way it works and the conditions of the environment in which it lives. In doing so they reveal the structure of functional interdependencies in the living world. That’s how design explanations are explanatory.

Summary: Explanation Without A Cause

Introduction

The topic of my dissertation is explanation in functional biology. Functional biology is that part of biology which is concerned with the life of individual organisms: the way they are built, the way they work and the way they behave. Functional biologists often explain why an organism is built the way it is, why it works the way it works, or why it behaves the way it does, by appeal to the functions of an item[1] or behaviour pattern.

An example is Schwenk's (1994) explanation of why snakes have forked tongues.[2] Schwenk explains this forking by appealing to the function of the tongue. He argues that the snake's tongue has an important function in following scent trails (of preys and mates). Snakes follow scent trails by comparing the intensities of chemical stimuli at two sides of the body. This allows them to detect the edges of a chemical trail and follow it with minimal deviation. The chemical stimuli are collected by means of the tongue and passed to a pair of chemoreceptors in the snout. In order to compare stimulus intensities at two points, the snakes must be able to sample chemicals at two points simultaneously. This is made possible by the forking. Hence, Schwenk explains the forked character of the tongue by showing that this form enables the tongue to fulfil a certain function (namely to collect chemical samples at two points simultaneously).

Another example is Krogh's (1941) explanation of the presence of a circulatory system of in many kinds of animals (vertebrates, annelids, crustaceans and so on).[3] One of the functions of the circulatory system is to circulate oxygen. Krogh uses the laws a physical-chemistry to show that an organism in which the distance between the outside and the inner organs exceeds 1 mm would not be viable if all oxygen transport was to be achieved by diffusion. Diffusion is not fast enough to meet the needs of the inner cells. Larger organisms need a system of convection in addition to diffusion. The circulatory system solves this problem by providing a system of convection. Other organisms employ other kinds of convection systems. Insects, for instance, transport oxygen by means of small tubes that circulate air and sponges and coelenterates trans-

---

[1] I use the term 'item' to refer to any part or substance of an organism (molecules, sub cellular organelles, cells, tissues, organs, organ systems and so on). Examples of items are chlorophyll molecules, chromosomes, the cell nucleus, membranes, white blood cells, hearts, limbs, fingers, the circulatory system and so on.

[2] Section 2.2.2, example 2.3.

[3] Section 4.2.3.

port oxygen by means of water currents. Hence, Krogh explains the presence of a circulatory system by showing that this systems fulfils a necessary function.

Biologist tend to call explanations such as those of Schwenk and Krogh ‘functional explanations’. Philosophers use the term ‘functional explanation’ also to refer to other kinds of explanations. To avoid confusion, I introduce a new term, namely ‘design explanation’, to refer to the kind of explanations exemplified above. Design explanations explain the presence, form or activity of an item or behavioural pattern (in a certain kind of organisms) by appealing to the function of that item or behavioural pattern (in that group of organisms).

To many people, functional explanations seem a little odd, at least at first sight. Their intuition tells them that explanations should show how the phenomenon to be explained is brought about by the explanatory facts. For example, an explanation of a solar eclipse tells us how an eclipse results from the moon sliding between the sun and the earth as a result of which the light from the sun cannot reach the earth. However, Schwenk’s explanation does not tell us how the tongue became forked. Neither does Krogh’s explanation show how larger organisms came to have a circulatory system.

In my dissertation, I develop a philosophical theory of design explanation. This theory addresses the question what design explanations add to our knowledge (in addition to the phenomena described in the explanation). For example, Schwenk describes the form of the tongue, he describes how snakes use their tongue and he relates how the latter phenomenon explains the first. What does this latter account add to the description of the phenomena described? Similarly, Krogh uses certain data about an organism’s need for oxygen and a well-known law of physical chemistry to show that an organism can never become large if oxygen transport is by means of diffusion alone. What does this calculation add to the data and the law?

Kinds of function

My account of design explanation starts with the observation that biologists use the term ‘function’ in a number of different ways. In the first part of chapter 2,[4] I distinguish four different uses of the term ‘function’, namely

(1) function as activity (function$_1$): what an item does or is capable of doing (its activities and capacities),
(2) function as causal role (function$_2$): the role of an item or behaviour pattern in maintaining a complex activity or capacity,
(3) function as survival value (function$_3$): the way in which a certain item or behaviour pattern contributes to the survival, reproduction or fitness of the organisms that have it,

---

[4] Section 2.2

(4) function as selected effect (function$_4$): the advantages of a certain trait for which it was selected in the past.

The first three kinds of function correspond to different meanings of the term ‘function’ as this term is used by functional biologists. The fourth kind of function is embraced by many philosophers (this use of the term function was introduced by the evolutionary biologist George Williams (1966) but this use of the term ‘function’ made no headway among biologists).

When biologists contrast ‘form’ and ‘function’ they typically use the term function in the sense of *function as activity* (function$_1$). Roughly spoken, ‘form’ refers to what an item is made of, the way it is built and the way it looks like and ‘function’ (in this sense) to what an item does or is capable of doing. Examples of activity descriptions are: ‘the glands in the mouth secrete saliva’, ‘the heart beats’, ‘the heart rate of normal human beings at rest is about 70 beats per minute’. Activity descriptions tell us what a certain item does, but they do not detail how this activity is important in a larger context. I shall use the term ‘character’ to refer to both the form and the activity characteristics of an item. Activity characteristics have no special position in design explanation.

The notions of function as causal role (function$_2$) and function as survival value (function$_3$) are, on the other hand, of special importance to understand design explanation. The failure to distinguish these two kinds of function is a main obstacle towards a theory of design explanation.

Attributions of *causal roles* concern the position of an item or behaviour in the way in which the execution of a certain task is organized.[5] An example of such an attribution can be found in Schwenk’s explanation of the forked tongue: ‘the function of the tongue of snakes in trail-following is to sample two points at one time’. This claim positions the tongue in the system that has the task to follow scent trails. It details the subtask of the tongue in that system.

Statements about *survival value* concern the presence or the character of a certain item or behaviour. Such statements compare the organism that interests us with another, hypothetical, organism in which the item or behaviour in question is absent or in which that item or behaviour has a different character. Such statements claim that in certain conditions the real organism is better off than the hypothetical one and detail why this is the case. An organism is better off than another organism if the fitness[6] of the first organism is higher than that of the second. An example of a statement about survival value is: ‘the forked character of the snake’s tongue is

---

[5] Computer scientists would say that the causal role is the logical position of an item in a system (in contrast to its physical position) that performs a certain task. I use the, admittedly tiresome, phrase ‘position in the way in which the execution of a certain task is organized’ to express the same idea. The term ‘organization’ is closer to the biologist’s language than the term ‘logical’.

[6] Fitness is a technical term in biology. It roughly means ‘the expected number of offspring’.

useful to the organisms that have it, because it enables them to follow scent trails more efficiently than they would do if their tongue was blunt'. This statement compares a real snake with a forked tongue with a hypothetical snake with a blunt tongue. It says that if the tongue has a causal role in trail-following, the real snake is better off than the hypothetical one, because that causal role can be performed with a forked one but not with a blunt one.[7]

Comparison between a real organism and a hypothetical one is usually called 'counterfactual comparison'. Functional biologists routinely apply counterfactual comparisons in design explanations. Philosophers are wary of this kind of comparison. I argue that this suspicion is undeserved. It results from a confusion of function as causal role and function as survival value. Once the distinction between these two kinds of function is properly drawn, there remain no objections to the use of counterfactual comparison to determine survival value.[8]

Explanation in functional biology.

In the second part of chapter 2,[9] I situate function attribution and design explanations in the practice of functional biology . I show by means of examples that functional biologists aim to answer seven types of questions. The topics of research and the products of enquiry are summarized in the table below.

Research in functional biology.

Topic of research	Products of enquiry
Form and activity characteristics	Descriptions of the structure and activity of organisms and their parts
Causal roles	Attributions of causal roles
Causes and underlying mechanisms	Physiological explanations
Survival value of performing certain tasks	Design explanations (of the presence of the item or behaviour that performs the causal role)
Survival value of having a certain character	Design explanations (of the character of the item or behaviour in question)
Ontogeny	Developmental explanations
Evolution	Evolutionary explanations

---

[7] Section 2.2.2, 2.2.3 and 6.2.

[8] Section 6.3.

[9] Section 2.3.

The foundation of explanation in functional biology is provided by accurate *descriptions* of organisms, parts and behaviours.

Attributions of causal roles provide the key to explanation in functional biology. I return to this issue after I have described the four kinds of explanations in which functional biologists are involved.

Physiological explanations detail the causes and underlying mechanisms of the phenomenon to be explained. In line with Cummins (1975, 1983) , I distinguish two kinds of physiological explanations:

(a) explanations that explain how a certain type of change in the state of an organism is brought about (transition explanations), and

(b) explanations that explain the properties of an item or behaviour in terms of underlying structures and mechanisms (property explanations).

Property explanation may involve two kinds of analysis:

(i) componential analysis: the part that has the property to be explained is analyzed into components,

(ii) property analysis: the property to be explained is analyzed into subproperties

A capacity explanation is a special kind of property explanation in which the capacity of an item or an organism to perform a certain task (i.e. causal role) is explained by appeal to the capacities of the parts of that item or organism to perform a series of subtasks which add up to the capacity to be explained. Capacity explanations in functional biology typically involve both componential and property analysis. Note that capacity explanations attribute causal roles to the parts.

Design explanations are concerned with the survival value of the presence or character of a certain item or behaviour. They explain why a certain item or behaviour is present in certain kinds of organisms or why that item or behaviour has the character it has in those kinds of organisms, by showing that in the conditions that apply to those organisms the trait in question is more useful some another conceivable trait.

Developmental explanations explain how a certain item or behaviour is brought about in the course of an individual’s history. These explanations are of the same kind as physiological transition explanations.

Evolutionary explanations explain how a certain item or behaviour was brought about in the course of the history of the lineage. Evolutionary processes include mutation, gene flow, recombination, selection and genetic drift. Selection explanations are a special kind of explanations that attempt to explain the presence or character of a certain item or behaviour by appeal to past selection.

Attributions of causal roles tell us how an item or behaviour is situated in the organism's organisation. They provide a handle by means of which functional biologists get a grip on their subject matter. Such attributions are used in at least three different kinds of explanation:

1) capacity explanations—as I noted above, these explanations appeal to the causal roles of the parts of an item or behaviour in explaining a capacity of that item or behaviour;
2) design explanations—survival value is typically assessed in relation to causal role;
3) selection explanations—evolutionary-historical explanations typically appeal to the efficiency with which a certain causal role is performed.

Philosophical theories of explanation

Philosophers of science have developed several theories of explanation. Two of these general theories are often applied in philosophical theories of functional explanation: the inferential theory (Hempel & Oppenheim 1948) and the causal theory (Salmon 1984) . Those theories are often seen as competing definitions of the notion of explanation. The inferential theory defines an explanation as a (deductive or inductive) argument in which a description of the phenomenon to be explained is inferred from a combination of descriptions of the laws of nature and the conditions that apply to the phenomenon in question. According to this theory the explanation of an eclipse (mentioned above) is explanatory because it infers that the conclusion that the light of the sun cannot reach the earth from a description of the positions of sun, moon and earth, and the laws about the propagation of light. The causal theory defines an explanation as an account that details the processes or mechanisms that bring about the phenomenon to be explained. I have already explained why the explanation of an eclipse is explanatory according to a causal account.

I am not interested in such exercises in conceptual analysis. However, the theories of explanation I mentioned above might also be seen as (possibly complementary) hypotheses about what scientists learn from the kind of reasoning they call 'explanation'. According to the inferential theory such reasoning shows that the phenomenon to be explained is to be expected in view of the explanatory facts. According to the causal theory such reasoning makes plain how the phenomenon to be explained hangs together with other phenomena in the world. This theory assumes that the different events in our world are brought about by means of causal interactions between causal processes. Explanations fit the phenomenon to be explained into this pattern of processes and interactions; they show us how different events are causally connected. From this point of view the philosophical theories of explanations are highly relevant to my topic.

Philosophical analyses of function and functional explanation

In chapter 4 to 7 I discuss a number of philosophical analyses which are presented, by their proponents, as theories of function or functional explanation. These analyses suffer from a number of problems:

1) they do not refer to real examples of explanations in functional biology to which their analysis applies,
2) they fail to take properly into account that biologists use the term ‘function’ in a number of different ways (most notably, they fail to distinguish between function as causal role and function as survival value),
3) they fail to take properly into account that there are different kinds of explanations that appeal to functions (many authors ignore design explanations or confuse them with capacity explanation or selection explanations),
4) they ignore the complex structure of design explanations (philosophers tend to think of functional explanations as one sentence function attributions in answer to a why-question).

As a result of these problems the reader is often left without any clue about the kind of explanation to which a certain analysis is supposed to apply.

In chapter 4 I discuss the now classical attempts of Hempel (1959) and Nagel (1961, 1977). These attempts employ the inferential theory of explanation. Design explanations pose a problem to the inferential theory because there are often different ways to fulfil a need or task. For example, blood circulation, air tubes en water transport provide different means to meet the need for a system of convection in larger organisms. Hence, from the fact that a certain organisms is able to fulfil a certain task or meet one may not infer the presence of a particular item, behaviour or structure. This problem is known as ‘the problem of functional equivalents’.

Hempel accepts the existence of functional equivalents and draws the conclusion that the kind of reasoning which is usually called ‘functional explanation’ is merely heuristic. It helps us to discover new phenomena but fails to explain them. I argue that Hempel’s account fails to account for many insights provided by design explanations.[10]

Nagel argues that there are no real functional equivalents. If both the function and the conditions in which the function is to be performed are specified in detail there remains only one way to perform that function. For example, given the way in which vertebrates are built, circulating oxygen by means of air or water currents is no option. I argue that this move does not work. In many cases one may only exclude functional equivalents by including into the explaining law the condition that the phenomenon to be explained is present. However, a statement of the type ‘all vertebrates that circulate blood, circulate blood’ is not a law of nature but a truism.

---

[10] Section 4.1

Furthermore, one of the insights provided by a design explanation is the insight that different systems meet the same need. If one does not allow functional equivalents one cannot account for this insight.[11]

Most (but not all) recent attempts to account for so-called functional explanations employ the causal theory of explanation. Proponents of the causal account have another problem with design explanation. Design explanations seem to explain the presence of a certain item, behaviour or trait by appealing to the fact that a certain task or need is fulfilled. However, the fulfilment of that task or need is an effect of the presence of the item / behaviour / trait, not a cause. Proponents of the causal theory respond in different ways. Some (e.g. Schaffner 1993: 362-410) maintain (just as Hempel did) that the scientific value of so-called ‘functional explanations’ is merely heuristic, not explanatory. Others try to show that this kind of reasoning, if properly viewed, really gives insight in the causes of the phenomenon to be explained. There are two kinds of approaches within the latter group: the disposition theory (Bigelow & Pargetter 1987) and the etiological theory.

The three main approaches to functions and functional explanations in contemporary philosophy are: the causal role theory (Cummins 1975, 1983), the survival value approach (for example Canfield 1964, Wimsatt 1972, Ruse 1973, Bigelow & Pargetter 1987, Horan 1989), and the etiological theory (for example Neander 1980, 1983, Millikan 1984, 1989b, Neander 1991a, Millikan 1993a).

In chapter 5 I discuss Cummins’s causal role theory. According to this theory function attributions describe the role of an item in maintaining a capacity of a system to which that item belongs. For example, to say that the heart has the function to propagate the blood is to say that propagating the blood is what the heart does that accounts for the organism's capacity to circulate the blood. I argue that this theory applies to causal roles. There are, however, other uses of the term ‘function’, which are left unanalyzed. Cummins argues that function attributions (i.e. attributions of causal roles) are not meant to explain the presence of the item to which the function is attributed. Instead, they are used to explain a capacity of a system of which that item is a part. I argue that Cummins’s account applies to capacity explanations. However, my main interest is in design explanations and Cummins’s account fails to provide insight in this kind of explanation.[12]

In chapter 6 I discuss the survival value approach. This approach identifies functions with effects that currently make a causal contribution to the survival and reproduction of an individual organism. For example, the function of the heart is to propagate the blood if that is how hearts currently contribute to the survival and reproduction of the organisms that have hearts. I

---

[11] Section 4.2

[12] It is unclear whether Cummins ignores, misrepresents or denies the existence of design explanations.

argue that this kind of definition confuses function as causal role and function as survival value. The proponents of the survival value approach differ in the way in which they account for the explanatory force of such function attributions. Horan (1989) employs the inferential theory of explanation. Bigelow & Pargetter (1987) favour a causal theory.

According to Horan, functional explanations show us how a certain trait is maintained in the population. I argue that this idea is most promising, but that Horan's elaboration of this idea is utterly confused.[13] There are other authors who have done a better job. Sober (1984) argues that appeals to overall fitness (i.e. function as survival value) are explanatory because they might show why a certain trait frequency distribution is maintained in the population. Reeve & Sherman (1993) suggest that appeals to adaptation explain why a trait remains prevalent in a population. Their notion of adaptation is very similar to my notion of function as survival value. I argue that this use of function as survival value accounts for part of the insights provided by appeals to survival value. However, it cannot be the whole story. Design explanations provide insights in the way in which certain kinds of individuals hang together, in addition to insights in population level processes. The explications of Sober and Reeve & Sherman do not account for these insights at individual level relations.

The disposition theory of Bigelow & Pargetter sees a function as a *disposition* to have an effect (that enhances the fitness of the individuals that produce that effect) rather than as that effect itself. The disposition explains the subsequent occurrence of the effect in the same way as the fragility explains its actual breaking. Mitchell (1995) argues convincingly that even if a disposition can be seen as the cause of survival (it might explain why an organism with a certain kind of item survives) it may not be seen as the cause of the presence of that kind of item (it does not explain why that kind of item is there (why it is present). Hence, the disposition theory does not provide insight in design explanations.

In chapter 7 I discuss the etiological theory. This is presently the dominant theory among philosophers of science with an interest in biology. According to this theory the function of an item / trait is to produce the effects for which it was selected in the past and which explain the item's / trait's current presence in the population. For example, hearts have the function to propagate the blood, if and only if propagating the blood is what hearts did that caused them to be favoured by past natural selection. This theory emphasizes that functions should not be seen as present effects of the item to which the function is attributed, but rather as past effects of past occurrences of that item in an ancestral population. Past effects, can be causes of present phenomena, of cause, and the etiological theory identifies the functions of an item with those past consequences that *were*, as a matter of fact, causally effective in the evolution or maintenance of

---

[13] Section 6.6

the item to which the function is attributed. I argue that this notion of function (function$_4$) is a philosophical fantasy that does not provide insight in design explanations.

I draw the conclusion that neither the inferential theory nor the causal theory provide an adequate account of the insights supplied by design explanations.

The structure of design explanation.

In chapter 8, I provide my own account of functional explanation. In the first part[14] I present my account of the structure of design explanations. According to this account a design explanation has two parts. It starts by specifying the causal role of the item or behaviour that interests us. Next, it discusses the survival value of the presence or character of that item or behaviour (on the basis of the preceding attribution of a causal role). In regard to the topic of explanation, I distinguish two variants: design explanations that explain the presence of the kind of item or behaviour to which the function is attributed (an example is Krogh’s explanation of the presence of a circulatory system in several groups of animals) and design explanations that explain the character of the kind of item or behaviour to which the function is attributed (an example is Schwenk’s explanation of the forked character of the snake’s tongue). Both kinds start with an attribution of a causal role but they differ in their second component. Design explanations of the presence of a certain item or behaviour in certain kinds of organisms point out why it is useful to perform this causal role in the conditions that apply to the organisms in question (for example, Krogh points out that the circulatory system meets the need for a system of convection created by the size of vertebrates). Design explanations of the character of a certain item or behaviour explain why given this causal role and given the conditions in which this role is to be performed the role is better performed in the way it is performed than in some other conceivable way (for example, Schwenk points out why the trail-following role of the snake’s tongue is better performed it the tongue is forked than if it is blunt).

Another distinction concerns the strength of the claim about survival value: requirement claims state that the trait to be explained is needed to survive and reproduce (examples are ‘a forked tongue is needed to sample two points simultaneously’ and ‘organisms of a certain size need a system of convection’), optimality claims state that the trait to be explained is the best among the alternatives considered (for example a deeply forked tongue is more useful than a lesser forked one). Many design explanations make several claims about survival value, some of them are optimality claims and others requirement claims.

Philosophers of science tend to ignore this complex structure of design explanations. They think of functional explanations as statements of the form ‘the function of *x* in organisms of taxon *t* is *y*’ in answer to a question of the form ‘why do organisms of taxon *t* have or perform

---

[14] Section 8.1.

x?’. It will be clear that this mistake is facilitated by the failure to distinguish between function as causal role and function as survival value. It will also be clear that this lack of insight in the complexity of design explanation is not conducive to understanding the explanatory force of this kind of reasoning.

The explanatory force of design explanations

In the second part of chapter 8[15] I discuss the explanatory force of design explanations and the nature of the relations with which design explanations are concerned. I argue that the explanatory force of design explanations must be sought in their ability to fit the presence and the character of the items and behaviours that interest us in the structure of functional interdependencies that exists between the different parts of an organism, its behaviour and the state of the environment in which it lives. The main difference between such relations of functional interdependence and causal relations is this: causal relations determine what is brought about (what happens), functional interdependencies determine what may exist (which combinations of traits of organisms and environmental states are viable). The most important relations of functional interdependence are ‘... is needed’ (for example, ‘vertebrates need a circulatory system’ and ‘... demands for ...’ (for example, the size of vertebrates demands for a circulatory system). Design explanations show us (i) how the properties of certain kinds of organisms and the states of the environment in which they live, pose a problem to the life of those organisms, and (ii) how those problems are solved in the organisms in question.

---

[15] Section 8.3 and 8.4.

Samenvatting: verklaren zonder oorzaken aan te wijzen

Dit boek gaat over verklaringen in de functionele biologie. Functionele biologie is dat deel van de biologie dat zich bezig houdt met individuele organismen: de manier waarop ze gebouwd zijn en de manier waarop ze functioneren. Functionele biologen doen vaak een beroep op de functie van een orgaan of gedragspatroon als ze de bouw, werking of levenswijze van een organisme verklaren.

Een voorbeeld is Schwenks (1994) verklaring van de vorm van de slangetong.[1] Het is algemeen bekend dat slangen een gespleten tong hebben. Schwenk verklaart deze gespletenheid met een beroep op de functie van die tong. Daartoe toont hij aan dat de slangetong een belangrijke functie heeft in het volgen van geursporen van prooien en paringspartners. Slangen gebruiken hun tong om te ruiken. Zij pikken met hun tong geursignalen op en brengen die naar een reukorgaantje in het verhemelte. Zij kunnen de richting van een geurspoor bepalen en voorkomen dat zij naar opzij afdwalen door de sterkte van de geur op verschillende punten van het spoor te vergelijken. Dit mechanisme werkt alleen als de geur van het spoor *tegelijkertijd* op twee punten verzameld wordt. Doordat de tong gespleten is, is een slang inderdaad in staat om op twee punten tegelijkertijd een geur te verzamelen. Schwenk verklaart de gespleten vorm van de tong dus door te laten zien dat deze vorm van belang is voor het vervullen van de functie van dat orgaan (nl. het verzamelen van geursporen op twee plekken op hetzelfde moment) in het volgen van geursporen.

Een ander voorbeeld is Kroghs (1941) verklaring van de aanwezigheid van een bloedvatstelsel in allerlei diergroepen (gewervelde dieren, wormen, kreeftachtigen etc.).[2] Een van de functies van het bloedvatstelsel is het transport van zuurstof. Krogh laat (op basis van natuur- en scheikundige theorieën) zien dat een beest waarin de afstand tussen de binnenkant van de organen en de buitenkant van het beest meer is dan ca. 1 mm niet kan bestaan zonder een actieve vorm van zuurstoftransport. Passief transport (diffusie) is niet snel genoeg om in de zuurstofbehoefte van de binnenste cellen te voorzien. Het bloedvatstelsel voorziet in een vorm van actief zuurstoftransport en lost dit probleem dus op. In sommige diergroepen is er een andere vorm van actief transport. Insekten transporteren zuurstof in holle buizen waardoor lucht stroomt. Sponzen en holtedieren (kwallen enzo) transporteren zuurstof in waterstromen. Krogh verklaart de aanwezigheid van een bloedvatstelsel dus door te laten zien dat dit orgaan een noodzakelijke functie vervult.

---

[1] Paragraaf 2.2.2, voorbeeld 2.3.

[2] Paragraaf 4.2.3.

Verklaringen zoals die van Schwenk en Krogh worden door biologen meestal ‘functionele verklaringen’ genoemd. Filosofen gebruiken de uitdrukking ‘functionele verklaring’ ook voor heel andere soorten verklaringen. Om verwarring te voorkomen heb ik een nieuwe term geïntroduceerd, namelijk ‘designverklaring’. Een designverklaring verklaart de aanwezigheid van een bepaald item[3] / gedragspatroon[4] (in een bepaalde groep organismen) of de manier waarop dat item / gedragspatroon gebouwd is (in die groep organismen). Zij doet dit aan de hand van de functie van dat item / gedragspatroon.

Veel mensen hebben het idee dat er met dit soort verklaringen iets raars aan de hand is. Zij zijn van mening dat verklaringen in de natuurwetenschappen moeten laten zien hoe het te verklaren verschijnsel ontstaat. In een goede verklaring van een zonsverduistering wordt bijvoorbeeld uitgelegd hoe die zonsverduistering tot stand komt: doordat de maan tussen de aarde en de zon schuift kan het licht van de zon de aarde niet meer bereiken. Schwenk laat echter (op het eerste gezicht in ieder geval) niet zien hoe de gespleten slangetong ontstaat. Evenmin laat Krogh zien hoe het bloedvatstelsel ontstaat.

In dit boek ontwikkel en verdedig ik een wetenschapsfilosofische theorie over designverklaringen. De centrale vraag die deze theorie moet beantwoorden is deze: ‘wat heb je aan een designverklaring?’. Meer precies geformuleerd wordt dit: ‘wat voegt een designverklaring toe aan de feiten die in de verklaring genoemd worden?’. Schwenk beschrijft de vorm van de slangetong (gespleten), hij beschrijft hoe de tong gebruikt wordt en vervolgens legt hij uit hoe het laatste verschijnsel het eerste verklaart. Wat voegt deze uiteenzetting toe aan de beschrijving van de verschijnselen? Wat leer je ervan? Krogh berekent aan de hand van gegevens omtrent de zuurstofbehoefte van organismen en een wet uit de fysische chemie dat een organisme zonder actief zuurstoftransport nooit groot kan worden. Wat voegt die berekening toe aan de combinatie van de gegevens en de natuurwet?

Er zijn in de wetenschapsfilosofie verschillende algemene theorieën van verklaring ontwikkeld. Twee van deze theorieën worden vaak toegepast op designverklaringen: de inferentiële theorie (Hempel & Oppenheim 1948) en de causale theorie (Salmon 1984) . Deze theorieën worden vaak gebracht als concurrerende definities van het begrip verklaring. Volgens de inferentiële theorie geeft iemand een verklaring als hij / zij een beschrijving van het te verklaren verschijnsel afleidt uit een combinatie van beschrijvingen van natuurwetten en beschrijvingen van

---

[3] Ik gebruik de term ‘item’ als algemene term voor een onderdeel van een organisme: orgaansysteem, orgaan, proces, weefsel, molecuul etc.. Voorbeelden van items zijn de tong en het bloedvatstelsel.

[4] Designverklaringen kunnen zowel betrekking hebben op items als op gedragingen. In het boek geef ik dan ook verschillende voorbeelden van designverklaringen in de ethologie. Om het kort te houden heb ik mij in de samenvatting beperkt tot twee voorbeelden uit de diermorfologie en laat ik de nuancering ‘item of gedragspatroon’ verder achterwege.

aanvangscondities. Volgens deze theorie is de eerder genoemde verklaring van een zonsverduistering een verklaring omdat uit een beschrijving van de posities van de zon, de maan en de aarde en de wetten van de verbreiding van licht af te leiden is dat het licht van de zon de aarde niet kan bereiken. Volgens de causale theorie geeft iemand een verklaring als hij / zij beschrijft hoe een bepaald verschijnsel tot stand komt. Ik heb al uitgelegd dat de eerder genoemde verklaring van een zonsverduistering volgens deze definitie een verklaring is.

Ik ben niet zo geïnteresseerd in dit soort oefeningen in conceptuele analyse. De hierboven genoemde algemene theorieën van verklaring laten zich echter ook interpreteren als elkaar mogelijk aanvullende opvattingen over wat je leert van het soort uiteenzetting dat door wetenschappers ‘verklaring’ genoemd wordt. Volgens de inferentiële theorie leer je van een verklaring dat het te verklaren verschijnsel te verwachten is (op grond van de natuurwetten en de aanvangscondities). Volgens de causale theorie leer je van een verklaring hoe het te verklaren verschijnsel samenhangt met andere verschijnselen in de wereld. Deze theorie poneert dat er tussen allerlei verschijnselen in de wereld een causale samenhang bestaat (de ene gebeurtenis brengt een andere voort). Een verklaring leert ons hoe het te verklaren verschijnsel in deze causale samenhang past. Zo geïnterpreteerd zijn deze theorieën van het grootste belang voor mijn vraagstelling.

In hoofdstuk 4 t/m 7 behandel ik een aantal filosofische analyses die door hun voorstanders als theorie over functie of functionele verklaring gepresenteerd worden. Al deze analyses lijden aan de volgende vier problemen: (1) er worden nauwelijks of geen concrete voorbeelden (met referenties naar publicaties in de biologie) gegeven van het soort verklaring waarop de analyse van toepassing zou zijn, (2) er wordt niet of te weinig onderkend dat biologen de term ‘functie’ op een aantal verschillende manieren gebruiken, (3) er wordt niet of te weinig onderkend dat er verschillende soorten verklaringen zijn waarin een beroep op functies gedaan wordt, (4) de complexe structuur van designverklaringen wordt over het hoofd gezien.

Het ontbreken van referenties naar biologische literatuur maakt het moeilijk om vast te stellen over wat voor soort functies en wat voor soort verklaring de betreffende analyse gaat.

De analyse van designverklaringen in de biologie wordt gecompliceerd doordat biologen de term ‘functie’ in meer dan één betekenis gebruiken.[5] Ik onderscheid drie betekenissen van de term ‘functie’ zoals die in de functionele biologie gebruikt wordt: (1) functie als activiteitskenmerk (functie$_1$), (2) functie als causale rol (functie$_2$), en (3) functie als overlevingswaarde (functie$_3$). Filosofen hebben de zaak extra ingewikkeld gemaakt door een vierde functiebegrip te introduceren: functie als een geselecteerd effect (functie$_4$).[6]

---

[5] Paragraaf 2.2

[6] Paragraaf 2.2.4 en 7.3.

Als biologen vorm en functie contrasteren, gebruiken ze de term ‘functie’ doorgaans in de eerste betekenis: functie als activiteitskenmerk (functie$_1$). Vormkenmerken betreffen het uiterlijk van een item, het materiaal waaruit het bestaat en de manier waarop het item gebouwd is (de structuur van dat item). Activiteitskenmerken betreffende de activiteit van dat item (wat doet het? hoe vaak? hoe snel? etc.). Voorbeelden van beschrijvingen van activiteitskenmerken zijn ‘de klieren in de mond scheiden speeksel uit’, ‘het hart klopt’ en ‘de hartslag van normale mensen in rust is ongeveer 70 slagen per minuut’. Activiteitsbeschrijvingen vertellen ons wat een bepaald item doet, zonder het belang van dat item voor een groter geheel daarin te betrekken. Vormkenmerken en activiteitskenmerken worden door biologen op dezelfde manier behandeld. De vorm- en activiteitskenmerken van een item tezamen noem ik het ‘karakter’ van dat item. Activiteitskenmerken hebben geen speciale positie in designverklaringen.

Functie als causale rol (functie$_2$) en functie als overlevingswaarde (functie$_3$) zijn wel van speciaal belang voor het begrijpen van designverklaringen. Het niet onderkennen van het onderscheid tussen deze twee is zelfs één van de belangrijkste hindernissen voor een goed begrip van dit soort verklaringen. Beweringen over de causale rol (functie$_2$) van een item betreffen de positie van dat item in de organisatie[7] van een systeem dat een bepaalde taak vervuld. Een voorbeeld van een toeschrijving van een causale rol (functie$_2$-toeschrijving) is te vinden in de eerder genoemde verklaring van de gespleten slangetong: ‘de functie van de tong van slangen bij het volgen van geursporen is het op twee plaatsen tegelijkertijd verzamelen van een geursignaal’. Deze bewering plaatst de tong in het systeem dat tot taak heeft geursporen te volgen en geeft aan dat de tong binnen dat systeem de deeltaak heeft om op twee plaatsen tegelijkertijd geursignalen te verzamelen. Beweringen over overlevingswaarde (functie$_3$) betreffen de aanwezigheid of het karakter van een bepaald item. Dergelijke beweringen maken een vergelijking tussen het organisme waarin we geïnteresseerd zijn en een ander, hypothetisch, organisme waarin het betreffende item afwezig is of een ander karakter heeft. De bewering geeft aan dat in bepaalde omstandigheden het echte organisme beter in staat is om te overleven dan het hypothetische organime zou zijn en waarom dat zo is. Een organisme is beter af dan een ander organisme als de fitness[8] van het eerste organisme groter is dan die van het tweede. Een bewering over overlevingswaarde berust doorgaans op een eerdere toeschrijving van een causale rol. Een voorbeeld van een dergelijke bewering is ‘het is nuttig voor een slang dat zijn tong gespleten is

---

[7] In de informatica zouden we zeggen dat een causale rol de logische positie in een systeem is (in contrast met de fysische positie). Hier druk ik hetzelfde idee uit door te zeggen dat een causale rol de positie in de organisatie van een systeem is. De term ‘organisatie’ sluit beter bij het taalgebruik van biologen aan dan de term ‘logisch’.

[8] Fitness is een technische term uit de biologie die zoiets betekent als het te verwachten aantal nakomelingen.

omdat hij op die manier in staat is geursporen te volgen'.[9] In deze bewering wordt een echte slang met een gespleten tong vergeleken met een hypothetische slang met een niet gespleten tong. De bewering stelt dat in de omstandigheid dat de tong een causale rol heeft bij het volgen van geursporen, de echte slang beter af is dan de hypothetische slang, omdat een gespleten tong deze causale rol kan vervullen, terwijl een niet gespleten tong dat niet kan.[10]

Het vergelijken van een bestaand organisme met een hypothetisch, niet-bestaand organisme wordt doorgaans 'counterfactual vergelijken' genoemd. Beweringen waarin een counterfactual vergelijking gemaakt wordt (bijvoorbeeld 'als een slang geen gespleten tong zou hebben, dan zou hij niet in staat zijn op twee punten tegelijkertijd geursporen te verzamelen'), heb ik 'functional counterfactuals' gedoopt. Biologen maken dit soort vergelijkingen routinematig. Filosofen zijn zeer argwanend tegenover dit soort vergelijkingen. Ik betoog dat deze argwaan voor een groot deel berust op een verwarring van causale rol en overlevingswaarde. Er is geen enkel bezwaar tegen counterfactual vergelijking voor het vaststellen van overlevingswaarde.[11]

Designverklaringen bestaan uit een toeschrijving van een causale rol (functie$_2$) aan het item waarin we geïnteresseerd zijn, gevolgd door een uiteenzetting betreffende overlevingswaarde op basis van deze toeschrijving. Er zijn twee vormen. In de eerste vorm wordt het karakter van een bepaald item verklaard, door te laten zien dat in de relevant omstandigheden het item zijn causale rol beter vervult als dat item het karakter heeft dat het heeft, dan wanneer het een ander karakter zou hebben. Een voorbeeld daarvan is de reeds vaker genoemde verklaring van de gespeten slangetong. In de tweede vorm wordt de aanwezigheid van een bepaald item verklaard door te laten zien dat dit item een essentiële causale rol vervult (d.w.z. dat het organisme beter af is als er een item is dat die causale rol vervult dan wanneer een dergelijk item ontbreekt).[12] Een voorbeeld daarvan is Kroghs verklaring van de aanwezigheid van een bloedvatstelsel.[13]

Filosofen zien deze complexe structuur van designverklaringen vaak over het hoofd. Zij vatten functionele verklaringen op als beweringen van de vorm 'de functie van x in organismen van taxon[14] t is y' in antwoord op een vraag van het type 'waarom hebben organismen van

---

[9] In de praktijk worden vaak verwarrende formuleringen gebruikt als 'de functie van de gespetenheid van de tong is het mogelijk maken van het volgen van geursporen' ('de function of the forked tongue is to enable trail-following').

[10] Paragraaf 2.2.2, 2.2.3 en 6.2.

[11] Paragraaf 6.3.

[12] Eerder heb ik deze tweede vorm van designverklaringen 'viability explanations' genoemd (Wouters 1995).

[13] Paragraaf 2.3.2 en 8.2.

[14] 'Taxon' is de naam die biologen gebruiken om een diergroep van een willekeurige rang aan te duiden, zoals fylum, familie, genus, species. Voorbeelden van taxa zijn: gewervelde dieren, zoogdieren, primaten, mensen.

taxon *t* een *x*?’. Het zal duidelijk zijn dat je deze fout eerder maakt als je geen onderscheid maakt tussen functie als causale rol en functie als overlevingswaarde.

Toeschrijvingen van een causale rol (functie$_2$-toeschrijvingen) worden toegepast in verschillende soorten verklaringen. In dit boek noem ik er drie: (1) capaciteitsverklaringen, (2) designverklaringen, (3) selectieverklaringen. Capaciteitsverklaringen verklaren hoe een item of een organisme een bepaalde taak kan uitvoeren door deze taak op te spitsen in een aantal deeltaken die elk door een onderdeel van dat item of van het organisme worden uitgevoerd.[15] Selectieverklaringen verklaren de aanwezigheid van een item met een bepaald karakter in de populatie met behulp van het mechanisme van natuurlijke selectie. Merk op dat er geen één op één relatie is tussen functiebegrippen en soorten als ‘functionele verklaringen’ gepresenteerde verklaringen.

De belangrijkste pogingen om vanuit een inferentiële theorie van verklaring het soort redenering dat als “functionele verklaring” aangeduid wordt te begrijpen zijn die van Hempel (1959) en die van Nagel (1961: 401-428, 1977). Voor een inferentiële theorie vormen designverklaringen een probleem omdat er vaak verschillende manieren zijn om in een taak of behoefte te voorzien. In de behoefte aan een actief transportsysteem kan bijvoorbeeld voorzien worden d.m.v. een bloedvatstelsel, maar ook door middel van luchtbuizen of een watertransportsysteem. Uit het feit dat een bepaald organisme in staat is in een bepaalde taak of behoefte te voorzien, kunnen we dus niet afleiden dat de te verklaren structuur / het te verklaren item te verwachten is.

Hempel trekt de conclusie dat de redeneringen die “functionele verklaringen” genoemd worden uitsluitend van belang zijn als hulpmiddel voor het zoeken naar verschijnselen. Dergelijke redeneringen helpen niet om de ontdekte verschijnselen te verklaren. Ik betoog dat Hempel ongelijk heeft omdat een designverklaring wel degelijk nieuwe inzichten levert in aanvulling op de in die verklaring beschreven verschijnselen.[16]

Nagel betoogt dat we het verschijnsel dat er meerdere manieren zijn om in een bepaalde taak of behoefte te voorzien kunnen uitsluiten door de voorwaarden waaronder in die taak of behoefte voorzien moet worden nauwkeurig te specificeren. Gezien de manier waarop gewervelde dieren gebouwd zijn, zou een luchtbuizen of watertransportsysteem niet werken. Ik betoog dat deze aanpak tekort schiet. In veel gevallen kan het bestaan van verschillende manieren om in een behoefte te voorzien alleen uitgesloten worden door in de verklarende natuurwet de voorwaarde op te nemen dat het te verklaren verschijnsel aanwezig is. Een bewering van het type ‘alle gewervelden dieren die een bloedvatstelsel hebben, hebben een bloedvatstelsel’ is echter geen natuurwet. Afgezien daarvan is het inzicht dat verschillende systemen in verschillende

---

[15] Deze analyse van capaciteitsverklaringen ontleen ik aan Cummins (1975, 1983).

[16] Paragraaf 4.1.

diergroepen in dezelfde behoefte voorzien nu juist één van de belangrijke inzichten die een designverklaring kan opleveren.[17]

Voor aanhangers van de causale theorie van verklaren vormen designverklaringen een probleem omdat ze de aanwezigheid of het karakter van een item lijken te verklaren met beroep op het feit dat ze in een behoefte van dat organisme voorzien. Het voorzien in een behoefte is echter geen oorzaak van de aanwezigheid van dat item / karaktertrek maar een effect. Aanhangers van de causale theorie hebben hier op twee manieren op gereageerd. Sommigen (bijvoorbeeld Schaffner 1993: 362-410) houden vol dat de redeneringen die biologen ‘functionele verklaringen’ noemen slechts waarde hebben bij het ontdekken van verschijnselen maar niet bij het verklaren daarvan. Anderen proberen te laten zien dat dergelijke redeneringen wel degelijk inzicht geven in de oorzaken van het te verklaren verschijnsel. Er zijn ruwweg twee manieren van aanpak binnen de laatste groep: de dispositietheorie en de etiologische theorie.

Een belangrijke poging om zogenaamde ‘functionele verklaringen’ te begrijpen vanuit de causale theorie is die van Cummins (1975). Ik betoog dat zijn analyse betrekking heeft op capaciteitsverklaringen en geen inzicht levert in de verklarende kracht van designverklaringen.[18]

De dispositietheorie (Bigelow & Pargetter 1987) vat een functie op als een effect van de aanwezigheid van een item of karaktertrek in een bepaald organisme dat de dispositie heeft bij te dragen aan de fitness van dat organisme. Mitchell (1995) laat op overtuigende wijze zien dat een dispositie misschien als de oorzaak van overleven opgevat kan worden (het kan verklaren waarom een organisme met een bepaald type item overleeft) maar niet als een oorzaak van het ontstaan van dat type item (het verklaart niet waarom een item van dat type ontstaat). De dispositietheorie levert dus geen inzicht in designverklaringen.

Een invloedrijke poging om zogenaamde ‘functionele verklaringen’ vanuit de causale theorie te begrijpen is de etiologische theorie (Neander 1980, 1983, Millikan 1984, 1989b, Neander 1991a, Millikan 1993a). Volgens deze theorie moet de functie van een item of karaktertrek in een bepaald organisme niet gezien worden als een deelverzameling van de effecten van dat item / karaktertrek in dat organisme, maar als een deelverzameling van de effecten die dat item / karaktertrek in het verleden had op de voortplanting van zijn dragers in een voorouderpopulatie. Effecten in het verleden kunnen natuurlijk wel gezien worden als oorzaken van verschijnselen in het heden en de etiologische theorie identificeert de functie van een item / karaktertrek met precies die effecten welke via het mechanisme van natuurlijke selectie bijdroegen tot de uitbreiding of handhaving van het betreffende item / karaktertrek in de populatie. Ik betoog dat dit begrip van functie als geselecteerd effect (functie$_4$) aan de fantasie van filosofen ontsproten is

---

[17] Paragraaf 4.2.

[18] Hoofdstuk 5.

en geen inzicht levert in designverklaringen in de functionele biologie (en evenmin in selectieverklaringen).[19]

Meer algemeen kan tegen iedere poging om designverklaringen op te vatten als selectieverklaringen worden ingebracht dat selectie een proces is dat zich op populatieniveau afspeelt terwijl designverklaringen een verband lijken te leggen tussen allerlei verschijnselen op individueel niveau. Schwenk legt een verband tussen de vorm en de taak van de tong in een bepaalde individu. Krogh legt een verband tussen de grootte van een organisme en de aanwezigheid van een bloedvatstelsel in dat organisme.

De conclusie is dat noch de inferentiële theorie noch de causale theorie ons voldoende inzicht geven in de verklarende kracht van designverklaringen.

Volgens mijn eigen theorie van designverklaringen moet de verklarende kracht van dergelijke verklaringen in eerste instantie gezocht worden in hun vermogen om de aanwezigheid en het karakter van een bepaald item in te passen in de structuur van functionele afhankelijkheden (“functional interdependencies”) die bestaat tussen de diverse onderdelen van een organisme, het gedrag van dat organisme en de toestand van de omgeving waarin het leeft. Het belangrijkste verschil tussen relaties van functionele afhankelijkheid en causale relaties is dit: causale relaties bepalen wat er ontstaat (wat er gebeurt), relaties van functionele afhankelijkheid bepalen wat kan bestaan (welke combinaties van karaktertrekken en omgevingsfactoren levensvatbaar zijn). De belangrijkste relaties van functionele afhankelijkheid zijn ‘ ... is nodig’ (bijvoorbeeld ‘gewervelde dieren hebben een bloedvatstelsel nodig’) en ‘... maakt nodig ...’ (bijvoorbeeld ‘de grootte van gewervelde dieren maakt dat zij een bloedvatstelstel nodig hebben’). Designverklaringen laten ons zien hoe bepaalde eigenschappen van een organisme en de toestand van de omgeving waarin het leeft een probleem vormen voor het bestaan van dergelijke organisme en hoe dit probleem in de betreffende organismen is opgelost.[20]

---

[19] Hoofdstuk 7.

[20] Paragraaf 8.3 en 8.4.

Dankwoord

Het was een lange en moeizame reis, die niet altijd even voorspoedig verliep. Velen hebben mij kortere of langere tijd vergezeld. Daarvoor wil ik hen bedanken.

Allereerst mijn begeleiders: Jaap van Brakel en Ton Derksen. Ik heb veel van jullie geleerd.

Jan Bergstra heeft, toen ik uren op mijn proefschrift zat te turen zonder een regel verder te komen en ik bovendien qua werk geen toekomst meer voor me zag, mij niet alleen aan werk geholpen (waardoor mijn toekomst er nu een stuk zonniger uitziet), maar mij er ook toe gezet het proefschrift toch maar af te maken.

De leden van de onderzoeksgroep ‘Rationaliteit en Naturalisme’ (Herman de Regt, Monica Meijsing, Chris Buskes, Pauwel Slurink, Menno Hulswit, Anthonie Meijers en Ton Derksen) wil ik bedanken voor de vele stimulerende discussies, de gezellige etentjes en de onderhoudende uitjes.

Cor van der Weele, Theo Kuipers, Wim van der Steen, Rob Visser, Hans Metz, Diedel Kornet en Rick Looijen dank ik voor hun commentaar op delen van mijn proefschrift of daar aan gerelateerde papers, en, meer algemeen voor hun inzet om de filosofie van de biologie in Nederland levend te houden. Ik betreur het ten zeerste dat dit vakgebeid in ons land steeds mee in de knel komt.

Ruth Millikan, Walter Bock, Anne Ruth Mackor en Cynthia MacDonald leverden nuttig commentaar op delen van mijn proefschrift of daar aan gerelateerde papers, waarvoor mijn dank.

Ik wil ook de deelnemers en begeleiders van de cursus “Phylogenie en Ecologie” van het (helaas) voormalige Instituut voor Theoretische Biologie in Leiden bedanken. Het heeft me goed gedaan via jullie nog enige contact met biologen te hebben. Ook het (korte) contact met Fred Janssen (didactiek van de biologie in Utrecht) betekende veel voor me. Het was goed te merken dat mijn werk zo bruikbaar voor je bleek.

Mijn huidige collega’s, met name Alex Sellink, Chris Verhoef, Paul Klint en Jan Bergstra, dank ik voor de nieuwe kansen en voor de tijd die ik kon vrijmaken om dit proefschrift af te ronden.

Dan zijn er de mensen met wie ik in Utrecht een kamer deelde: eerst Rob van Gerwen. later Desiree van Hattum. Jullie waren mij ook in het persoonlijke vlak tot steun.

Mijn borgtochgenoten, Cor van der Weele en Jeanne Peijnenburg, bedank ik voor de aanmoediging, voor de wijze raad en voor het aanhoren van mijn klachten. Ik stel het zeer op prijs dat jullie mij nu als paranimf terzijde staan.

Freek Wiedijk en Karst Koymans dank ik voor hun hulp met de computer, voor de gezelligheid en aanmoediging en voor de apparatuur die jullie mij lieten testen. Albert Visser wil ik bedanken voor de gezelligheid, voor de aanmoediging, en voor de samenwerking tijdens de cursus ‘taalfilosofie’. Het seminar ‘Niet-Euclidische Meetkunde’ was voor mij een uitzonderlijk genoegen.

Tenslottte wil ik Josje Lodder bedanken voor de aanmoediging, voor de werkschema’s, voor het lezen en verbeteren van allerlei stukken tekst, voor het verdragen van mijn vermoeidheid en slechte humeur en, last but not least, voor de discussies over de logische eigenschappen van de functionele verbanden die in dit proefschrift centraal staan. Ik hoop dat we er nog een keer een artikel over zullen schrijven.

Ondanks de nieuwe perspectieven in de informatica doet het me pijn te moeten constateren dat dit proefschrift tevens als zwanezang geldt.

References

Abramson, N.I. (1993), Evolutionary Trends in the Dentition of True Lemmings (Lemmini, Cricetidae, Rodentia): functional-adaptive analysis. *Journal of Zoology (London)* **230:** 687-699.

Achinstein, P. (1977), Function Statements. *Philosophy of Science* **44:** 314-367.

Alcock, J. (1989), *Animal Behavior. An Evolutionary Approach.* Sunderland: Sinauer.

Amundson, R. & Lauder, G.V. (1994), Function Without Purpose: The Uses of Causal Role Function in Evolutionary Biology. *Biology and Philosophy* **9:** 443-469.

Antonovics, J. (1987), The Evolutionary Dys-Synthesis: Which Bottles for Which Wine. *The American Naturalist* **129:** 321-331.

Appel, T. (1987), *The Cuvier-Geoffroy Debate: French Biology in the Decades before Darwin*. New York: Oxford University Press.

Ayala, F.J. (1970), Teleological Explanations in Evolutionary Biology. *Philosophy of Science* **37:** 1-15.

Balinsky, B.I. (1975), *An Introduction to Embryology*. Philadelphia: W.B. Saunders Company.

Baublys, K.K. (1975), Comments On Some Recent Analyses of Functional Statements in Biology. *Philosophy of Science* **42:** 496-486.

Beckner, M. (1959), *The Biological Way of Thought.* New York: Columbia University Press.

Bennett, J. (1976), *Linguistic Behaviour*. Cambridge: Cambridge University Press.

Bigelow, J. & Pargetter, R. (1987), Functions. *The Journal of Philosophy* **84:** 181-196.

Bock, W.J. (1961), Salivary Glands in the Gray Jays (*Perisoreus*). *Auk* **78:** 355-365.

Bock, W.J. (1980), The Definition and Recognition of Biological Adaptation. *American Zoologist* **20:** 217-227.

Bock, W.J. (1990), From Biologische Anatomie to Ecomorphology. *Netherlands Journal of Zoology* **40:** 254-277.

Bock, W.J. & von Wahlert, G. (1965), Adaptation and the Form-Function Complex. *Evolution* **19:** 269-299.

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

Brandon, R.N. (1978), Adaptation and Evolutionary Theory. *Studies in History and Philosophy of Science* **9:** 181-206.

Brandon, R.N. (1981), Biological teleology: Questions and Explanations. *Studies in History and Philosophy of Science* **12:** 91-105.

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

Brooks, D.R. & McLennan, D.A. (1991), *Phylogeny, Ecology and Behavior.* Chicago: The University of Chicago Press.

Burnet, M. (1962), The Thymus Gland. *Scientific American*, **207** (5): 50-57.

Canfield, J. (1964), Teleological Explanation in Biology. *British Journal for the Philosophy of Science* **14:** 285-295.

Canfield, J. (1965), Teleological Explanation in Biology: A Reply. *British Journal for the Philosophy of Science*

Canfield, J. (1966), Introduction. In: Canfield, J. (ed.), *Purpose in Nature*; Englewood Cliffs: Prentice Hall, p.

Catchpole, C.K. (1979), *Vocal Communication in Birds.* London: Edward Arnold.

Coddington, J.A. (1988), Cladistic Tests of Adaptational Hypothesis. *Cladistics* **4:** 3-22.

Cohen, G.A. (1978), *Marx's Theory of History: A Defense.* Princeton: Princeton University Press.

Coleman, W. (1964), *Georges Cuvier, Zoologist: A Study in the history of evolution theory*. Cambridge: Harvard University Press.

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

Cummins, R. (1983), *The Nature of Psychological Explanation.* Cambridge: The MIT Press.

Cuvier, G. (1817), *Le Règne Animal distribué d'apres son organisation.* Paris:

Darwin, C. (1859), *On the Origin of Species.* London: John Murray.

Darwin, C. (1886), *The Descent of Man, and Selection in Relation to Sex.* New York: D. Appleton.

Dobzhansky, T. (1956), What is an adaptive trait? *American Naturalist* **90:** 33-347.

Dobzhansky, T. (1968), Adaptedness and Fitness. In: Lewontin, R.C. (ed.), *Population Biology and Evolution*; Syracuse: Syracuse University Press, p. 109-121.

Dow, D.D. (1965), The Role of Saliva in Food Storage by the Gray Jay (*Perisoreus canadensis*). *Auk* **82:** 139-154.

Dullemeijer, P. (1974), *Concepts and Approaches in Animal Morphology*. Assen: Van Gorcum.

Eliott, S., Tait, N.N. & Briscoe, D.A. (1993), A Pheromonal Function for the Crural Glands of the Onychophoran *Cephalofovea tomahmontis* (Onychophora: Peripatosidae). *Journal of Zoology (London)* **231:** 1-9.

Endler, J.A. (1986), *Natural Selection in the Wild*. Princeton: Princeton University Press.

Fisher, R. (1930), *The Genetic Theory of Natural Selection*.

Fraassen, B.v. (1980), *The Scientific Image*. Oxford: Oxford University Press.

Frankfurt, H.G. & Poole, B. (1966), Functional Analyses in Biology. *British Journal for the Philosophy of Science* **17:** 69-72.

Futuyma (1986), *Evolutionary Biology.* Sunderland: Sinaur.

Gans, C. (1962), The tongue protrusion mechanism in *Rana catesbeiana. American Zoologist* **2:** 524-.

Gegenbaur, C. (1870), *Grundzüge der Vergleichenden Anatomie*. Leipzig: Wilhelm Engelmann.

Godfrey-Smith, P. (1994), A Modern History Theory of Functions. *Nous* **28:** 344-362.

Goldsmith, T.H. (1990), Optimization, Constraint, and History in the Evolution of the Eyes. *The Quaterly Review of Biology* **65:** 281-322.

Gotmark, F., Winkler, D.W. & Anderson, M. (1986), Flock-feeding on fish schools increases individual success in gulls. *Nature* **319:** 589-591.

Gould, S.J. (1980), Sociobiology and the Theory of Natural Selection. In: Barlow, N. & Silverberg, A. (ed.), *Sociobiology: Beyond Nature/Nurture?*; Boulder: Westview, p. 59-64.

Gould, S.J. & Lewontin, R.C. (1979), The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London* **B 205:** 581-598.

Gould, S.J. & Vrba, E.S. (1982), Exaptation — A Missing Term in the Science of Form. *Paeobiology* **8:** 4-15.

Griffiths, P.E. (1993), Functional Analysis and Proper Functions. *British Journal for the Philosophy of Science* **44:** 409-422.

Habibi, K., Thouless, C.R. & Lindsay, N. (1993), Comparative behaviour of sand and mountain gazelles. *Journal of Zoology (London)* **229:** 41-53.

Harvey, P.H. & Pagel, M.D. (1991), *The Comparative Method in Evolutionary Biology*. Oxford: Oxford University Press.

Harvey, W. (1628), *Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus*.

Hempel, C.G. (1959), The Logic of Functional Analysis. In: Gross, L. (ed.), *Symposium on Sociological Theory*; New York: Harper and Row, p. 271-287.

Hempel, C.G. (1965), *Aspects of Scientific Explanation*. New York: Free Press.

Hempel, C.G. & Oppenheim, P. (1948), Studies in the Logic of Explanation. *Philosophy of Science* **15:** 135-175.

Horan, B.L. (1989), Functional Explanations in Sociobiology. *Biology and Philosophy* **4:** 131-158, 205-228.

Hull, D.L. (1974), *Philosophy of Biological Science*. Englewood Cliffs: Prentice-Hall.

Hutchinson, J.M.C., McNamara, J.M. & Cuthill, I.C. (1993), Song, Sexual Selection, Starvation and Strategic Handicaps. *Animal Behaviour* **45:** 1153-1177.

Huxley, H.E. (1965), The Mechanism of Muscular Contraction. *Scientific American*, **213** (6): 18-27.

Johansen, K. (1977), Respiration and Circulation. In: Kluge, A.B. (ed.), *Chordate Structure and Function*; New York: Macmillan, p. 306-391.

Kant, I. (1790), *Kritik der Urteilskraft*. Berlin: Lagarde und Friederich.

Keeton, W.T. & Gould, J.L. (1993), *Biological Science*. New York: Norton.

Kettlewell, B. (1959), Darwin’s Missing Evidence. *Scientific American*, **200** (3): 48-53.

Kettlewell, B. (1973), *The Evolution of Melanism*. Oxford: Clarendon Press.

Kluge, A.G. (ed.) (1977),*Chordate Structure and Function*. New York: Macmillan.

Krimbas, C.B. (1984), On Adaptation, Neo-Darwinian Tautology, and Population Fitness. In: Hecht, M.K. & Wallace, B. (ed.), *Evolutionary Biology*; New York: Plenum, p. 1-57.

Krogh, A. (1941), *The Comparative Physiology of Respiratory Mechanisms*. Philadelphia: University of Pennsylvania Press.

Kuipers, T.A.F. (1985), The Logic of Intentional Explanation. *Communication and Cognition* **18:** 221-242.

Kuipers, T.A.F. (1986a), Explanation by Specification. *Logique et Analyse* **116:** 509-521.

Kuipers, T.A.F. (1986b), The Logic of Functional Explanation in Biology. In: *The Tasks of Contemporary Philosophy. Proc. 10th Wittgenstein Symp., Kirchberch/Wechsel,1985*; Vienna: Hölder-Pitcher-Tempoky, p. 110-114.

Kuipers, T.A.F. (1996), Explanation by Intentional, Functional, and Causal Specification. *Poznan Studies in the Philosophy of Science and Humanities* **47:** 209-236.

Kuipers, T.A.F. & Wisniewski, A. (1994), An Erotetic Approach to Explanation by Specification. *Erkenntnis* **40:** 377-402.

Lack, D. (1947), *Darwin's Finches: an essay on the general biological theory of evolution.* Cambridge: Cambridge University Press.

Langman, R.E. (1989), *The immune system*. San Diego: Academic Press.

Lankester, E.R. (1870), On the use of the term Homology in Modern Zoology, and the distinction between Homogenetic and Homoplastic agreements. *Ann. Mag. Nat. Hist* **4:** 35-43.

Lauder, G.V. (1986), Homology, Analogy and the Evolution of Behavior. In: Nitecki, M.H. & Kitchell, J.A. (ed.), *Evolution of Animal Behavior*; New York: Oxford University Press, p. 9-40.

Lauder, G.V. (1990), Functional Morphology and Systematics: Studying Functional Patterns in an Historical Context. *Annual Review of Ecology and Systematics* **21:** 317-340.

Lauder, G.V. & Liem, K.F. (1989), The role of historical factors in the evolution of complex organismal function. In: Wake, D.B. & Roth, G. (ed.), *Complex Organismal Functions: Integration and Evolution of Vertebrates*; Chichester: John Wiley and Sons, p.

Lehman, H. (1965a), Functional Explanation in Biology. *Philosophy of Science* **32:** 1-20.

Lehman, H.S. (1965b), Teleological Explanation in Biology. *British Journal for the Philosophy of Science* **15:** 327.

Lenoir, T. (1982), *The Strategy of Life*. Dordrecht: Reidel.

Levin, M.E. (1976), On The Ascription of Functions to Objects, with Special Reference to Inference in Archaeology. *Philosophy of the Social Sciences* **6:** 227-234.

Lewontin, R. (1981), The Organism as the Subject and Object of Evolution. *Scientia* **118:** 65-82.

Lewontin, R. (1983), Darwin's Revolution. *New York Review of Books* **30:** 21-72.

Mackie, J.L. (1965), Causes and Conditions. *American Philosophical Quarterly* **2:** 245-264.

Mackie, J.L. (1974), *The Cement of the Universe.* Oxford: Clarendon.

Malinowski, B. (1948), *Magic, Science and Religion*. New York:

Marler, P. (1975), On the Origin of Speech from Animal Sounds. In: Kavanagh, J.F. & Cutting, J. (ed.), *The Role of Speech in Language*; Cambridge: MIT Press, p. 11-37.

Mauzerall, D. (1977), Porphyrins, Chlorophyll, and Photosynthesis. In: Threbst, A. & Avron, M. (ed.), *Photosynthesis I*; Berlin: Springer Verlag, p. 117-124.

Mayr, E. (1959), Typological vs. Population Thinking. In: *Evolution and anthropology: a centennial appraisal*; Washington: The Anthropological Society of Washington, p. 409-412.

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

Mayr, E. (1976), *Evolution and the Diversity of Life*. Cambridge: Harvard University Press.

Mayr, E. (1982), *The Growth of Biological Thought*. Cambridge: Harvard University Press.

Mayr, E. (1997), *This is Biology: The Science of the Living World*. Cambridge: Harvard University Press.

McClamrock, R. (1993), Functional Analysis and Etiology. *Erkentniss* **38:** 249-260.

McNeill Alexander, R. (1975), *The Chordates*. Cambridge: Cambridge University Press.

McNeill Alexander, R. (1979), *The Invertebrates*. Cambridge: Cambridge University Press.

McNeill Alexander, R. (1988), Why Mammals Gallop. *American Zoologist* **28:** 237-245.

Merton, R.K. (1957), Manifest and Latent Functions. In: *Social Theory and Social Structure*; New York: The Free Press, revised and enlarged edition, p. 19-84.

Miller, J.F.A.P. (1961), Immunological Function of the Thymus. *The Lancet,* 30 September 1961, p. 748-749.

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

Millikan, R.G. (1986), Thoughts Without Laws; Cognitive Science With Content. *Philosophical Review* **95:** 47-80.

Millikan, R.G. (1989a), An Ambiguity in the Notion “Function”. *Biology and Philosophy* **4:** 172-176.

Millikan, R.G. (1989b), In Defense of Proper Functions. *Philosophy of Science* **56:** 288-302.

Millikan, R.G. (1993a), Propensities, Exaptations, and the Brain. In: *White Queen Psychology and Other Essays for Alice*; Cambridge: The MIT Press, p. 31-50.

Millikan, R.G. (1993b), *White Queen Psychology and Other Essays for Alice*. Cambridge: MIT Press.

Mills, S. & Beaty, J. (1979), The Propensity Interpretation of Fitness. *Philosophy of Science* **49:** 268-273.

Mitchell, S.D. (1989), The Causal Background of Functional Explanation. *International Studies in the Philosophy of Science* **3:** 213-230.

Mitchell, S.D. (1993), Dispositions or Etiologies? A comment on Bigelow and Pargetter. *The Journal of Philosophy* **90:** 249-259.

Mitchell, S.D. (1995), Function, Fitness and Disposition. *Biology and Philosophy* **10:** 39-54.

Nagel, E. (1961), *The Structure of Science*. London: Routledge & Kegan Paul.

Nagel, E. (1977), Teleology Revisited. *The Journal of Philosophy* **74:** 261-301.

Neander, K. (1980), Teleology in Biology. *Unpublished*

Neander, K. (1983), *Abnormal Psychobiology.* La Trobe University. Unpublished PhD thesis.

Neander, K. (1991a), Function as Selected Effects: The Conceptual Analyst’s Defense. *Philosophy of Science* **58:** 168-184.

Neander, K. (1991b), The Teleological Notion of ‘Function’. *Australasian Journal of Philosophy* **69:** 454-468.

Nissen, L. (1970), Canfield’s functional translation schema. *British Journal for the Philosophy of Science* **21:** 193-201.

Owen, R. (1843), *Lectures on Invertebrate Animals.*

Pap, A. (1962), *An Introduction to The Philosophy of Science*. New York: The Free Press of Glencoe.

Pranger, R. (1990), Towards a Pluralistic Concept of Function. *Acta Biotheoretica* **38:** 63-71.

Reeve, H.K. & Sherman, P.W. (1993), Adaptation and the Goals of Evolutionary Research. *The Quarterly Review of Biology* **68:** 1-32.

Rescher, N. (1964), *Hypothetical Reasoning.* Amsterdam: North Holland.

Resnik, D.B. (1995), Functional Language and Biological Discovery. *Journal for General Philosophy of Science* **26:** 119-134.

Reznek, L. (1987), *The Nature of Disease*. London: Routledge & Kegan Paul.

Ridley, M. (1993), *Evolution.* Cambridge: Blackwell.

Rimmer, D. (1998), Ultimate Interface. *NewScientist,* 14 November 1998, p. Inside Science number 15.

Robinson, J.D. (1986), Reduction, Explanation, and the Quests of Biological Research. *Philosophy of Science* **53:** 33-353.

Romer, A.S. & Parsons, T.S. (1986), *The Vertebrate Body*. Philadelphia: Saunders.

Ruse, M. (1973), *The Philosophy of Biology*. London: Hutchinson.

Russell-Hunter, W.D. (1968), *A Biology of Lower Invertebrates.* London: MacMillan.

Salmon, W.C. (1978), Why Ask, ‘Why?’? *Proceedings and Addresses of the American Philosophical Association* **51:** 683-705.

Salmon, W.C. (1984), *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

Salmon, W.C. (1989), *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.

Schaffner, K.F. (1993), *Discovery and Explanation in Biology and Medicine*. Chicago: The University of Chicago Press.

Schwenk, K. (1994), Why Snakes Have Forked Tongues. *Science* **263:** 1573-1577.

Seely, G.R. (1977), Chlorophyll in Model Systems: Clues to the Role of Chlorophyll in Photosynthesis. In: Barber, J. (ed.), *Primary Processes of Photosynthesis*; Amsterdam: Elsevier, p. 1-50.

Simon, H.A. (1957), *Models of Man: social and rational*. New York:

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

Sober, E. (1993), *Philosophy of Biology*. Oxford: Oxford University Press.

Sorabji, R. (1964), Function. *Philosophical Quarterly* **14:** 289-302.

Stelt, A.v.d. (1968), *Spiermechanica en Myotoombouw bij Vissen*. Amsterdam: Veen.

Thorpe, W.H. (1958), The learning of song patterns by birds, with especial reference to the song of the chaffinch *Fringilla coelebs. Ibis* **100:** 535-570.

Tinbergen, N. (1963), On Aims and Methods of Ethology. *Zeitschrift für Tierpsychologie* **20:** 410-433.

Tinbergen, N. (1976), *Animal Behaviour*. S.l: Time Life International.

Tinbergen, N., Broekhuysen, G.J., Feekes, F., Houghton, J.C.W., Kruuk, H. & Szulc, E. (1962), Egg Shell Removal by the Black Headed Gull, *Larus ridibundus* L. *Behaviour* **19:** 74-117.

Ward, P. & Zahavi, A. (1973), The importance of certain assemblages of birds as “information-centres” for food finding. *Ibis* **115:** 517–534.

Wehner, R. & Bernard, G.D. (1993), Photoreceptor Twist: A solution to the false-color problem. *Proceedings of the National Academy of the Sciences (USA)* **90:** 4132-4135.

Wehner, R., Bernard, G.D. & Geiger, E. (1975), *Journal of Comparative Physiology* **A 134:** 109-112.

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

Wimsatt, W.C. (1972), Teleology and the Logical Structure of Function Statements. *Studies in History and Philosophy of Science* **3:** 1-80.

Woodfield, A. (1976), *Teleology*. Cambridge: Cambridge University Press.

Wouters, A.G. (1995), Viability Explanation. *Biology and Philosophy* **10:** 435-457.

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

Wright, L. (1976), *Teleological Explanations: An Etiological Analysis of Goals and Functions.* Berkeley: University of California Press.

Zweers, G.A. (1979), Explanation of Structure by Optimization and Systemization. *Netherlands Journal of Zoology* **29:** 418-440.