Biological Functions and Natural Selection: A Reappraisal

Marc Artiga
Universitat de València

(forthcoming in *European Journal for Philosophy of Science*)

Abstract

The goal of this essay is to assess the Selected-Effects Etiological Theory of biological function, according to which a trait has a function F if and only if it has been selected for F. First, I argue that this approach should be understood as describing the paradigm case of functions, rather than as establishing necessary and sufficient conditions for function possession. I contend that, interpreted in this way, the selected-effects approach can explain two central properties of functions and can satisfactorily address some recent counterexamples. This reading, however, shows that there is only a partial overlap between biological functions and selected effects, so the former cannot be reduced to the latter. Finally, I maintain that this result is still compatible with a naturalistic theory of function that appeals to some evolutionary process.

1 Introduction

Biological functions have two striking features.[1] On the one hand, an appeal to functions seems to have some special explanatory value. More precisely, it is often suggested that a trait's function partly explains why that trait

[1]This paper is about *biological* functions. For brevity, I will sometimes speak of 'functions', but please keep in mind that I always mean 'biological function'. For a complementary account of artifact function, see Artiga (submitted).

exists: eyelashes' having the function of protecting the eye explains why many animals have them. Yet how can an item's effect explain its own existence? And why do functions rather than any other effects occupy this explanatorily privileged position? On the other hand, functions set a normative standard: an entity is *supposed* to perform its function. If it fails to have this effect, it is said to *be dysfunctional* or to *malfunction*. But what is the nature of these norms? How do they originate? Is this fact compatible with naturalism? Let us call these explananda 'EXPLANATION' and 'NORMATIVITY'. A central goal of a theory of function is to provide an account that enables us to better understand these properties (Wouters, 2005; Garson, 2016: 3-7).

One of the most popular ways of approaching this challenge is the so-called 'Selected-Effects Etiological Theory', according to which the notion of a function can be analyzed in terms of selection for. The goal of this paper is to assess this approach. First of all, I will argue that the Selected-Effects Etiological Theory should be understood as specifying the paradigm functional effect. This interpretation allows the theory to explain why functions exhibit EXPLANATION and NORMATIVITY and shows how this approach can address some recent counterexamples. The second main claim of the paper is that adopting this perspective has significant consequences: on the interpretation I will develop, there is only a partial overlap between functions and effects being selected for. Whereas paradigm cases of function largely overlap with selected effects, marginal instances fail to do so. A consequence of this merely partial overlap is that functions cannot straightforwardly be reduced to selected effects, as traditionally assumed. Finally, I will contend that this result is still compatible with the truth of some sort of etiological theory (to which the Selected-Effects Etiological Theory might be a good approximation).

The paper is structured as follows. In section 2 I will briefly present the Selected-Effects Etiological Theory of function. In section 3 I will argue that, if we conceive of this theory as defining a paradigm from which actual instances can differ in various respects, this approach can account for EXPLANATION and NORMATIVITY, and can also address some puzzling counterexamples. In section 4 I will argue that a similar analysis can be provided for 'selection for'. From these results, however, I will argue, in section 5, that there is only a partial overlap between functions and selected effects. Nonetheless, I will also contend that this claim is compatible with the truth of some form of etiological theory. I will conclude with a short summary of the main results.

2 The Selected-Effects Etiological Theory

According to the Selected-Effects Etiological Theory, the notion of function can be explained in terms of selection for:

> **(T1)** A trait T has the function F iff T has been selected for F.[2]

Obviously, T1 only provides a satisfactory theory of biological function if 'selection for' can be appropriately defined. The standard strategy resorts to evolutionary theory: roughly, a trait T has been selected for F if and only if F is the effect of T that causally explains why T was favored by natural selection (Millikan, 1984; Neander, 1991a, 1991b; Griffiths, 1993; Godfrey-Smith, 1994).[3] More precisely, 'selection for' is usually defined along the following lines:

> **(T2)** T has been selected for F iff F was an effect of T and the following conditions hold:
>
> 1. *Heredity*: Individuals reproduce and offspring tend to resemble their parents.
> 2. *Variation*: Individuals differed from one another in whether they possessed a trait T that performed F.
> 3. *Differential Reproduction*: Individuals that possessed a trait T that performed F were more successful at reproducing than individuals that did not possess a trait T that performed F.
> 4. *Causation*: F causally contributed (positively) to reproduction.

This analysis can be divided into two main parts. On the one hand, the first three conditions are supposed to establish that T's performance of F is

[2]Selected-effect etiological theorists usually hold that for a trait to have a function it must have *recently* been selected for (Griffiths, 1993; Godfrey-Smith, 1994; Millikan, 2002: 123). Since this modification is largely irrelevant for the purposes of this paper and would unnecessarily complicate some of the claims and arguments I will develop below, in this paper I will simply assume that the relevant past is the recent one.

[3]'T' refers to a type. Tokens have the function F in virtue of being instances of type T (see, for instance, Neander, 1991b: 460).

the result of natural selection. Following the classical description of natural selection, this involves a population in which there is some variation among its members, which is in part heritable, and distinct variants leave different numbers of offspring (e.g. Lewontin, 1985: 76; Ridley, 2004: 74). For a trait to be selected for F, however, more than this is required, since it is a causal notion and traits might be favored by natural selection even if they do not have any causal impact on the process of evolution. A standard illustration of this idea is provided by pleiotropy, which occurs when a single genetic component influences two or more unrelated phenotypes, F_1 and F_2. If F_1 has some positive effect, natural selection can lead both F_1 and F_2 to fixation within a population, even if F_2 has no causal impact on fitness. Thus, for a trait T to be selected for F a fourth condition should be added: F must be an effect of T that causally contributed to success (Sober, 1984: 97-102).

Let us suppose for a moment that both T1 and T2 hold. From these two claims, it trivially follows that the notion of function can also be analyzed in terms of heredity, variation, differential reproduction, and causation. That is:

> **(T3)** T has the function F iff F was an effect of T and the following conditions hold:
>
> 1. *Heredity*: Individuals reproduce and offspring tend to resemble their parents.
> 2. *Variation*: Individuals differed from one another in whether they possessed a trait T that they performed F.
> 3. *Differential Reproduction*: Individuals that possessed a trait T that performed F were more successful at reproducing than individuals that did not possess a trait T that performed F.
> 4. *Causation*: F causally contributed (positively) to reproduction.

The conjunction of T1 and T2 describes the Selected-Effects Etiological Theory (or, for short, 'SEE Theory') and T3 is a logical consequence of this approach. As Neander (1991b: 459) suggests, "in my view, the central element of the etiological approach should be seen as the simple idea that a function of a trait is the effect for which that trait was selected." Likewise,

Millikan (1993: 40) writes that "a trait's biological 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" [emphasis in the original]. The rest of the paper is devoted to analyzing this simple but powerful idea.

3 Assessing T3

I will begin with a discussion of T3. Before addressing this claim, however, it is important to be clear about the kind of project we are engaged in. In principle, one could use T3 for different purposes. Neander (1991a), for instance, holds that T3 provides a conceptual analysis of the concept *function*. Here, I will follow some other researchers in understanding T3 as a theoretical analysis: we seek to understand better a particular entity that scientists aim at describing, rather than examining the ordinary user's conception of a term (Millikan, 1989; 2002: 115; Papineau, 2001; Schulte, forthcoming; Shea, 2018). As will be made clear below, the project is not primarily to capture our intuitions, but to specify the features or dimensions that are explanatorily relevant for an effect to qualify as a function. Thus, T3 is supposed to spell out the nature of functional effects.

A satisfactory theoretical analysis needs to meet two main desiderata. Firstly, T3 can only provide an accurate account of function if entities that satisfy T3 are coextensional with entities that have functions, i.e. clear cases of function should correspond to effects that fully comply with T3, whereas entities that to some extent fail to fulfill some of its conditions might either qualify as marginal cases or fail to possess functions at all.

Showing that fulfilling T3 is coextensional with having a function, however, does not suffice to vindicate a theoretical analysis, since the features listed in T3 could simply correlate with the presence of functions, without specifying its nature. In order to argue for the stronger claim, one should also show that the conditions listed in T3 partly explain the properties associated with functions (Schulte, forthcoming).[4] In other words, in the same way that being H_2O partly explains why water is transparent, liquid, has a

[4]As Millikan (1993: 17) suggests, "The various properties, the various analogies, which are influential in leading us to speak of quite diverse categories of items as having "functions" are properties or analogies that are characteristically *accounted for* by the fact that these items have coincident *proper functions*."

certain electrical conductance and so on, the fact that functions satisfy T3 should partly explain the central features associated with functions, such as EXPLANATION and NORMATIVITY. Can an etiological theory meet these challenges?

I think it definitely can. In what follows, I will argue that T3 (or a similar apprach that I will discuss) can satisfactorily explain why functions exhibit EXPLANATION and NORMATIVITY and why effects that fit this analysis are coextensional with functions. Nonetheless, I will draw some unexpected consequences from how these features are accommodated.

3.1 Accounting for EXPLANATION and NORMATIVITY

In this section I would like to show that T3 (or a slightly modified version of T3) can explain why functions set a normative standard (i.e. NORMATIVITY) and why they possess a special explanatory role (i.e. EXPLANATION). To accomplish this goal, three preliminary conceptual points need to be made.

First, some of the conditions from T3 might be relevant to accommodating some aspects of function, and others might be relevant for explaining other features. I think this is a general consequence of how kinds explain their properties. Suppose an entity E has two properties, F1 and F2, and imagine that, according to our best theory, typically an instance of E exists whenever two conditions hold, C1 and C2. In some cases C1 might be more relevant for explaining one of the features (e.g. F1) than C2. For instance, suppose that politically left-wing individuals (E) tend to vote for socialist or social democratic parties (F1) and recycle more than right-wingers (F2) and suppose further that according to our best theory of what it is to be a left-winger, it involves being egalitarian (C1) and environmentalist (C2)(Huber and Inglehart, 1995; Krange et al., forthcoming; McCright and Dunlap, 2011; Neumayer, 2004). If we then seek to explain why left-wing individuals recycle more than other groups, we would probably mention their being environmentalists, rather than their being egalitarian.

The second point is that these conditions can be satisfied in various degrees. An entity that clearly satisfies all conditions for being an E will be called 'paradigmatic' (Godfrey-Smith, 2009). The notion of 'paradigm' does not primarily aim to describe the most intuitive cases, but those that play some special explanatory role (recall that our goal is to provide a theoretical rather than a conceptual analysis). Paradigm instances are the central cases that clearly possess those properties that are most explanatory of a kind.

As a result, if a paradigm case of E possesses C1, and C1 explains why E has feature F1, one should expect a correlation between partial satisfaction of C1 and partly possessing F1. All things being equal, if a left-winger only partly defends environmentalism, one should expect him to recycle to a lesser extent than a paradigmatic left-winger.

Finally, in some cases not satisfying a particular condition entails not qualifying as an instance of E, whereas not fully complying with other conditions means qualifying as a marginal example of E. Although paradigmatic left-wingers support environmentalism, one of them could oppose it and still qualify as a left-winger — although in that case, she would not be paradigmatic. In contrast, a left-winger has to be an egalitarian (to some extent, at least). Someone who completely rejects egalitarianism cannot be classified as left-wing. The reason, of course, is that egalitarianism is much more central than environmentalism to the distinction between left-wing and right-wing political ideologies (Huber and Inglehart, 1995). For our purposes, what is essential is to realize that failing to satisfy some conditions disqualifies you from being a left-winger, whereas failing to comply with others could just make you a non-paradigmatic left-winger. Only some conditions must be fulfilled in order to be an E, while all of them must be satisfied to qualify as a paradigm E.[5]

Having clarified these distinctions, let us again focus our attention on T3. Functions possess two puzzling features, EXPLANATION and NORMATIVITY. Since, according to T3, functions are effects that comply with *Variation*, *Heredity*, *Differential Reproduction* and *Causation*, the fact that an effect satisfies these requirements should explain why it possesses these two puzzling features. Yet it is much easier to claim that effects that satify T3 exhibit EXPLANATION and NORMATIVITY than giving reasons for this statement. How can we justify this claim?

The strategy I will follow is to consider each of the conditions included in T3 separately and show how they contribute to explain these two puzzling properties of functions. For instance, I will show that the fact that a trait's effect E in a population P satisfies *heredity* contributes to explain why this

[5]For a biological example illustrating the same idea, think about recent definitions of *organism*, which seek to specify the key aspects that are central for being a clear case of organismality such as integration, cohesiveness, genetic uniformity, etc. Queller and Strassmann (2009), for instance, define what they call the 'paradigm organism' as an entity with high cooperation and low conflict among its parts and classify entities along these dimensions, in the same way we will do in section 3.2.

effect sets a normative standard (i.e. it exhibits NORMATIVITY). In a sense, the strategy is to vindicate the idea that effects that comply with T3 have a specially normative and explanatory role by decomposing T3 and showing how each of the different conditions contributes to account for these two aspects.

Furthermore, assessing to what extent each condition of T3 (*heredity*, *variation*, etc..) contributes to accounting for the two puzzling features of function (EXPLANATION and NORMATIVITY) will be important later on: we will consider some philosophers (e.g. Buller, Garson) who hold that satisfying some conditions of T3 is not required for a trait to have a function. One way of responding to these challenges is to show that each of these conditions makes an important contribution to the explanation of why functions play their distinctive normative and explanatory role. To do that, however, we need to discuss the contribution that some conditions makes on its own to account for NORMATIVITY and EXPLANATION.

3.1.1 Explanation

Most philosophers accept that functions have an important explanatory role to play, and many think this consists in the fact that an item's function partly explains why it exists. This is what we have called 'EXPLANATION'. But how can a trait's effect explain its own existence? According to common wisdom, to solve this puzzle we should realize that we are not trying to explain how the effect of a particular trait token at a time t_n explains why this very same token exists at t_n. Rather, the suggestion is that the effect of previous items of the same type at a previous time t_{n-x} explains the current existence of items of this type at t_n[6] (Millikan, 1984; Neander, 1991b; Griffiths, 1993).

At this point, it is also important to distinguish between two ways of interpreting EXPLANATION. The goal of explaining why a type of trait exists at a time t_n can be understood either as suggesting that functions explain why traits of this type *persisted* in a population or why they were *created* in the first place. Following Neander (1995), we will call the former 'persistence' and the latter 'origin' explanations. In the context of functions, the teleological aspect is usually understood in the former sense (Griffiths, 1993; Godfrey-Smith, 1994; Cummins, 2002: 163).[7] Thus, in this debate, EXPLANATION

[6]Recall that the function of a token depends on its belonging to a certain type. See footnote 3.

[7]For a very good reason: it is hard to see how functions could provide origin explana-

should be interpreted as requiring a persistence explanation (Wouters, 2005: 129).

We can now reformulate our initial question: which of the conditions from T3 are relevant for accommodating the fact that functions provide persistence explanations of trait types? Certainly, it is decisive that the trait token belongs to a lineage and that it is a reproduction of a past item: after all, the whole point of the standard strategy rests on showing that the ancestor's effects explain why present tokens of that trait exist. Thus, satisfying *Heredity* is central to T3's ability to account for EXPLANATION. Similarly, the fact that performing F causally contributed to reproduction is essential for grounding this explanatory role: functions are usually assumed to provide causal explanations, so if F is explanatory at all, it must have causally contributed to the trait's persistence. Consequently, fulfilling *Causation* is also very important.

The role of *Variation*, on the other hand, is much less clear. It is not obvious, for instance, why the fact that there was a significant amount of variation (or, perhaps, variation at all, as we will see in 3.2) should be important for accommodating EXPLANATION, i.e. the fact that functions explain why trait types exist. Imagine that there is a population in which all individuals possess a trait T that perform F (i.e. there is no variation) and in which F causally and positively contributes to the survival and reproduction of members of the population. In this case, it seems that T's performance of F could still provide a persistence explanation of T (Shea, 2018: 64).

One way of finding a place for *Variation* in accounting for EXPLANATION (but probably not the only one) appeals to robustness. Birch (2012: 299) argues that teleological language is often used for processes that 'robustly converge towards the same end-point despite variation of initial conditions'. He maintains that this is true even of chemical and physical processes: if almost any initial point would have led to the same result E, scientists are inclined to say that the process exists 'in order to' produce E. Hence, teleological language is employed in providing *robust-process explanations*, i.e. explanations that not only describe what happens in the actual world, but

tions. For one thing, that would require that the effect of a trait existed before the very trait evolved. For another, the standard way of accommodating EXPLANATION I sketched above suggests that the fact that present hearts pump blood is explained by some members *of the same type* pumping blood in the past and this explanans presupposes that hearts already existed. Notice that the claim that functions provide persistence explanations is fully compatible with natural selection underpinning origin explanations (see section 4).

also track what occurs in a range of possible worlds in which the initial conditions differ (see also Jackson and Pettit, 1990; Sterelny, 1995). Relying on this analysis, here is a way of finding a place for variation in our account of EXPLANATION: if, within a population, the process leading to the actual distribution of traits T performing F lacked significant variation in the recent past, then the process might have been highly unrobust, in the sense that a different trait with some slightly different effect F* could have easily done much better and wiped out Ts performing F from the population. In contrast, if, in the recent past, the population included diverse traits with various kinds of effects, then the process leading to the current presence of traits performing F is likely to be more robust. Thus, the existence of variation in a population supports the idea that performing F provides a robust-process explanation of the persistence of a certain trait within this population.[8]

If these arguments are on the right track, then some of the conditions of T3 can explain why functions have a special explanatory role, i.e. EXPLANATION.[9] Let us turn now to the second compelling feature of functions: NORMATIVITY.

3.1.2 Normativity

Kidneys are supposed to filter wastes from blood, and those which fail are said to *malfunction* or *function incorrectly*. This is the aspect we labeled 'NORMATIVITY'. Which of the conditions in T3 helps to explain this property? On the one hand, NORMATIVITY can be accommodated only if the conditions for having a function allow for malfunction, i.e. only if what a trait is supposed to do is not straightforwardly defined in terms of what it currently does. Accordingly, the fact that functions are defined by historical properties is central to T3's ability to account for this normative aspect.

[8]This is not to say that the teleological dimension can be fully analyzed in terms of robustness, or that this is the only way to connect *Variation* and EXPLANATION; my modest claim is just that this is a possible avenue for arguing that *Variation* (as well as *Differential Reproduction*) is a defining feature of functions that contributes to accounting for EXPLANATION.

[9]Of course, here I do not aim to provide an exhaustive analysis, since that would require a much longer and more detailed discussion from than the one I can offer. For instance, I do not examine *Differential Reproduction*, although it interestingly depends on the existence of *Variation*. To keep things manageable, I concentrate on the claims that are required to justify my main point.

This is true as far as it goes, but it might not go far enough: the fact that etiological theories base function attribution on history shows that a trait can have a function that it does not fulfill, but it is not obvious why this implies that this is an effect that the trait *should* have, in such a way that that there is something *wrong* if it does not have this effect.[10]

So why are functional attributions normative? It is often suggested that the fact that functions satisfy EXPLANATION can also ground this normative dimension (Price, 1995; Wouters, 2005: 129). On this view, the very idea that performing F explains why the trait exists suffices for accommodating the fact that the trait T should perform F. If a population contains individuals with T, and they persisted because they performed F, then this provides one reason for thinking that items of this type are supposed to do F; after all, F is the reason they are there. Thus, according to this perspective, the very same conditions and processes that account for the existence of the trait would suffice for grounding a weak form of normativity. To the extent that this is true, if the conditions in T3 explain EXPLANATION, then they also account for NORMATIVITY.

Still, some people might resist the idea that merely accounting for persistence is enough for grounding a norm of performance. It seems that a process can account for the maintenance or even spread of a certain phenomenon without generating any norm that distinguishes proper from improper effects (see section 3.2 for some suggestive examples).

There is, however, a different (and complementary) way of accounting for the normativity of functions, although it requires a slight modification of T3. If we think of paradigmatic cases of functions, these effects not only help explain why future tokens of the same type exist; they do so by contributing to the fitness of a larger system (e.g. an organism). Hearts' pumping blood at some time t_{n-x} helps explain why hearts will pump blood at t_n by contributing to the survival and reproduction of a living system. If a trait T assists a large device in doing some further thing by doing F, then this need seems to set a norm which T might fail to meet. Indeed, the idea that functions are contributions to the maintenance and reproduction of larger systems is an assumption of some SEE theorists (e.g. Godfrey-Smith, 1994; Griffiths, 1993; Price, 2001), and this is a powerful motivation for many alternative

[10] *Wrongness* should be understood here in a very weak sense, of course. Nothing like moral normativity is involved. Still, it is hard to see why the mere fact that function depends on past performance suffices for even a weak form of normativity.

accounts of functions as well (Boorse, 2002; Canfield, 1963; Cummins, 1975; Mossio et al., 2009; Sorabji, 1964; Woodfield, 1976). Thus, if to the conditions for having a function we add that this effect should have contributed in the recent past to the fitness of a larger system (or perhaps: to the survival and reproduction of a living being), the resulting account would be in a better position to account for the normativity of functions. It could account for a stronger sense in which a trait is supposed to perform its function, and ground the idea that if it fails to have this effect, it is dysfunctional or functions incorrectly. We will see later that some alleged counterexamples to T3 seem to point in the same direction.

Crucially, notice that this contribution to the fitness of a larger system is a past effect, rather than a current event. Hence, this is a version of an etiological theory, insofar as it suggests that "being preceded by the right kind of history is *sufficient* to set the norms that determine purposiveness" (Millikan, 1993: 26). This account differs from classical approaches that define functions as contributions to fitness or to the goals of a larger system, since these accounts typically reject any appeal to history in defining functions (Canfield, 1963; Boorse, 2002; Maley and Piccinini, 2018; but see Garson, forthcoming).[11]

Summing up, the two main striking properties of functions can be accounted for by the claims included in T3, although it might need to be supplemented with the further condition that the effect must have contributed to the fitness of a larger system. Henceforth, I will use 'T3+' to refer to T3 plus contributing to the fitness of a larger system and I will assume that this is the right way to go. All in all, it provides a compelling argument in favor of the idea that T3+ captures some of the central aspects of functions. Furthermore, note that these aspects can be satisfied to various degrees, so instances that clearly comply with all of them should count as paradigmatic, whereas those that do not meet this conditions (or meet them poorly – see

[11] Although a fair assessment of these approaches would require a more detailed analysis than what I can offer here, notice that if a trait's function depends on its current contribution to the organism's fitness (or on the fact that members of the same type currently contribute to fitness), then a trait's function fails to explain the existence or maintenance of a trait (so it does not fulfill EXPLANATION). Furthermore, as many philosophers have pointed out, this account has difficulties in dealing with malfunction: it is impossible, for instance, for all present traits to malfunction, because if all of them fail to contribute to fitness, then there is no basis for saying that they have a function in the first place (Neander, 1991a).

below) should qualify as marginal.

Of course, there are other motivations for T3 and T3+ besides the fact that they account for EXPLANATION and NORMATIVITY. One important argument is that by appealing to such features as heredity, one can avoid some of the counterexamples that were presented against early etiological theories such as Wright's (1973). For instance, according to Millikan (1993: 34), a break in the gas hose that is sustained by leaking gas that overcomes the man that has to fix it does not have a function because it is not the result of a reproductive process (this counterexample is due to Boorse, 1976: 72). In any case, the goal of this essay is not to consider all properties of functions, or to explore all possible ways of justifying two or three aspects, but to explore a particular way of understanding and defending the etiological theory of function (section 3) and build an argument against identifying functions with effects being selected for by natural selection (section 5).

3.2 Coextension

In the previous section I argued that the etiological theory of function can account for the two main puzzling features of functions. The goal of this section is to consider a second line of thought in favor of T3: assessing whether satisfying T3 is coextensional with having a function.[12] On the one hand, as many supporters of the SEE Theory have extensively shown, paradigm examples of functions (e.g. the heart's pumping blood) seem to fit T3 relatively well. Nonetheless, things become much harder (and more interesting) once we reflect on less central cases.

More precisely, in this section I would like to discuss, in some detail, different examples that have been provided in the literature of functional entities in which some conditions of T3 are only partly satisfied or, in some cases, not met at all. I will focus on this particular set of examples because they will help to motivate the two main claims of the paper. On the one hand, that T3 actually defines the paradigm case of function from which particular cases can differ in various respects. On the other, that there is a

[12]Since in this paper we are not only addressing the question of whether an entity is classified as a function or not, but also whether it counts as a paradigm or a marginal case, the notion of 'coextension' is similarly fine-grained: in the sense intended here, property A is coextensional with property B if and only if central cases of A are central cases of B and marginal cases of A are marginal cases of B. I want to thank a reviewer for pressing me on this point

partial overlap between functional effects and effects being selected for.

At this point it is worth keeping in mind that, as I argued in section 3, not fulfilling some of the conditions in T3 might have two sorts of consequences. In some cases, failing to satisfy a particular condition for being an E implies being a marginal instance of E, whereas on other occasions it entails not qualifying as an E at all. A left-winger who rejects environmentalism is a non-paradigmatic example of a left-winger, but a non-egalitarian just cannot be a left-winger. I will show that this distinction is important in the context of alleged counterexamples to *Variation*, *Heredity* and *Contribution to the Fitness of a Larger System*.

The first set of cases purporting to challenge T3 that I would like to discuss was presented by Buller (1998), who argued that variation is not required for a trait to acquire a function (see also Boorse, 2002; Abrams, 2005; Shea 2018). Lewens (2004: 92) illustrates this idea with a simple example: suppose there is a population of moths with bright orange wings that reside for most of the time on green leaves. Because of this mismatch, birds easily detect and eat them. However, at some point a chemical plant is build nearby, which releases a thick orange smoke that leaves a residue on the leaves. After this event, the orange months are well camouflaged, and as a result their predation risk decreases significantly. Suppose now that some time passes by, the population of orange moths increases, and for some reason no new variant emerges: during a certain period there are no moths in the population with a different color.[13] Intuitively, the function of the moth's wing color is to camouflage them, but this is not a result that T3 can get, since according to it, a trait T can acquire a function F only if performing F has been favored over traits that have other effects. Some people think this result is unsatisfactory: if orange wings contributed to fitness in the recent past, and this fact explains why orange moths have done better than they did before, then this should be enough for attributing a function to it, even if no variants were present. As a result, Buller maintains that *Variation* in T3 is not a necessary condition and puts forward his 'weak etiological theory', which "does not define the function of a trait in terms of selection for it, and thus does not make it a necessary condition of T's having a function that there was variation in T" (Buller, 1998: 508; 2002).[14]

[13]See Kreamer (2014) for a more realistic example.

[14]Schwartz (1999, 2002) develops a similar account (which he labels the 'Continuing Usefulness Account'), which requires variation at some point in the evolutionary past (but

Now, I doubt that this example really challenges T3. Rather, it seems to point to a case in which *Variation* is not satisfied, resulting in a marginal case of functionality. The reasoning developed in the previous section underpins this claim: as we argued there, if the process that causally accounts for the presence of F within a population does not involve variation, it might provide a much less robust explanation of F. In particular, if slight variations of trait T had existed in the recent past, then they could have easily done much better than F (a moth with slighlty darker orange wings, for instance, could have camouflaged much better). The fact that a population contains variation with respect to F contributes to robustly account for the distinctive explanatory role of function attributions (i.e. EXPLANATION), so a population that fails to contain variation cannot produce a paradigmatic functional trait, but only a marginal instance of function. Paradigm cases of function attributions provide more robust explanations of the persistence of a certain trait, and this aspect is probably missing from any process that lacks *Variation*.

A second objection to T3 has been presented by Garson (2011, 2012, 2017), who argues that a process of reproduction (presupposed by *Heredity* in T3) might not be required for a trait to have a function (see also Shea, 2018). His main argument focuses on neuronal selection, which is a process that underlies some forms of learning. For instance, in synapse selection, some synapses engage in a sort of competitive process that results in some of them being pruned (usually, those that become active less frequently). Garson argues that the synapses that survive seem to acquire a function (and he defends this position with some plausible examples) and maintains that this process mimics natural selection, with a single crucial difference: the synapse that is selected is not a reproduction of a previous one, but is only retained. As a consequence, Garson presents a disjunctive account of function, according to which a trait has a function F iff F contributed either to differential reproduction or to differential retention. Since the SEE Theory requires reproduction in order for a trait to have a function, Garson concludes that it should be modified (see also Millikan, 1993: 47; Neander, 2017: 1148; Dussault and Bouchard, 2017; Garson and Papineau, 2019).

Again, I think that Garson points to some interesting examples, but I disagree with the conclusions he draws from them. Synapse selection fails to fully satisfy *Heredity*, so it is a marginal case of function (in particular,

not recently).

it fails to fully satisfy the reproductive aspect[15]. Paradigm cases of function satisfy EXPLANATION, i.e. provide persistence explanations of a *type* of trait. Synapse selection can only provide a persistence explanation of a *token* trait. This is not the kind of explanation delivered by paradigm cases, so there are good reasons for classifying it as a marginal instance. For instance, if a token trait has a function in virtue of doing F, then malfunction is severely limited- it is not possible for a trait token to have always failed to perform F, for example. In contrast, when the function is attributed in virtue of belonging to a type, a token could have always been unable to F (Millikan, 1989). Similarly, scoring low on *Heredity* opens the door to a more liberal attribution of functions that can let it some processes that (to say the least) are not paradigmatic of functionality, as the classical example of a gas leakage is supposed to illustrate (for a discussion of this point, see Garson, 2017: 535-539). As a result, there are reasons for thinking that the examples provided by Garson are marginal cases, which can be accommodated by T3 and T3+. Therefore, the disjunctive definition of function he puts forward is probably unnecessary.

While the previous two challenges question whether T3 provides necessary conditions for having a function, others have pointed out that this criterion might not be sufficient. Godfrey-Smith (1994) and Manning (1997) discuss segregation distortion genes, which seem to be selected for F but do not have F as one of their functions. These genes disrupt normal gene division: by inducing sperm carrying the rival chromosome to self-destruct as they are formed, they manage to produce a higher frequency of themselves in the surviving chromosomes after meiosis. Disrupting meiosis is something that segregation distortion genes do that explains their survival and reproduction, so according to T3 this should be their function. Intuitively, however, they do not seem to have any function, so Godfrey-Smith (1994: 349) suggests extending T3 by adding that "the explanation of the selection of the functionally characterized structure must go via a positive contribution to the fitness of the larger system" (see also Price, 1995).

This proposal fits our previous suggestion that a definition of function should include the fact that the effect contributes to the fitness of a larger system. Adding this condition not only helps to account for the normative

[15]There are different ways of failing to comply with *Heredity*. A population in which there is low parent-offspring ressemblance, for example, would also score low on that dimension for a different reason.

aspect (as we saw above), but can also be used to exclude these counterexamples. Moreover, it correctly predicts that we do not usually attribute functions to whole organisms (unless we consider them to be parts of a larger system, such as an ecological system – Godfrey-Smith, 1994) and explains why cancer cells and clay crystals lack functions (Price, 1995: 152; Cf. Bedau, 1991).

These results lead to a remarkable conclusion: some of the recent counterexamples raised against T3 can be accommodated if we adopt T3+ and distinguish between paradigm and marginal cases. More precisely, these allegedly problematic cases should be considered marginal instances (or no instances at all) of function due their failing to fully comply with some conditions that are relevant for its possession. Furthermore, in the previous section we provided some motivations for these conditions, so we gave independent reasons for thinking of Buller's and Garson's examples as marginal and Godfrey-Smith's cases as not functional at all.[16]

3.3 Spatial Representation

To impose some order on the complex picture that results from this discussion, I think it will be useful to use Godfrey-Smith's (2009) spatial representation as a heuristic device (see also Millikan, 2017: 11-26). This tool is helpful because it represents in a neat way complex relationships. It also illustrates the role of gradients and the differences and similarities between particular examples along different dimensions (this will be important in my argument later on). Furthermore, it helps us think about new cases by displaying unexplored areas and suggest new questions for future research.

In a nutshell, the suggestion is to build a three-dimensional space, in which each axis corresponds to a particular aspect that is relevant for a given concept. Ideally, one should employ a higher-dimensional space, in which every feature that is relevant for possessing a function is included, but given the obvious limitations only three dimensions at a time will be considered here.

Since we are interested in analyzing functions, I suggest concentrating on *Heredity*, *Variation* and *Contribution to a larger system* (for reasons that will

[16]To dispel any remaining doubt about considering Buller's and Garson's examples to be marginal, notice that this category also admits of degrees: an item can be marginal with respect to some dimension without being marginal with respect to others. In other words: an item can qualify as a marginal case of marginality.

become clear in section 5). The idea, then, is to imagine a three-dimensional space in which each dimension corresponds to one of these features: the x axis corresponds to variation, the y axis to heredity and the z axis to contribution to a larger system (as represented in figure 1).[17] Within this space, we position the effects according to the values they obtain on each of these dimensions. For example, T's effect F scores highly on the axis *Contribution to a larger system* iff individuals possessing T recently contributed to the fitness of a larger system by F-ing. Similarly, it obtains a high score on *Heredity* iff individuals performing F reproduce and there is strong parent-offspring resemblance, i.e. the state of the parent is predictive to a great extent of the state of the offspring. Finally, T's effect F scores highly on *Variation* if this effect recently had to compete against other effects. If we focus on the main examples discussed so far in this paper, I think the result would look like figure 1. This figure graphically summarizes the main results of this section.

As can be observed in figure 1, the heart's pumping blood (which, following most philosophers, I take to be the paradigmatic example of a functional effect) would score highly in all three dimensions, since hearts recently contributed to the survival of organisms possessing them, are inherited and do better than hearts that fail to pump blood (or do so less efficiently). In contrast, Lewens' orange moths and synapse selection involve processes that lack variation and reproduction, respectively, and for this reason I have argued that they only qualify as marginal cases of function. Finally, segregation distortion genes are inherited and compete against genes that do not have these effects, but fail to positively contribute to the organism's fitness; thus, as Godfrey-Smith and Manning claim, they probably lack functions. Again, notice that failing to possess some properties may lead an effect to qualify as marginal, whereas lacking others entails not having a function at all.[18]

[17] Let me stress that Godfrey-Smith (2009) does not draw this figure. Here I am combining the results of our previous discussion with his suggestions about how to spatially represent concepts. Some significant differences between our approaches remain; for instance, he primarily focuses on Darwinian populations (and, secondarily, on Darwinian individuals- see Godfrey-Smith, 2009: 40), whereas my focus here is on traits and their effects.

[18] I mainly considered effects that occupy corners, but some intermediate cases could also be described. For instance, a population in which there is little variation with respect to a trait, would occupy a position near $\langle 0.5, 1, 1\rangle$. Part of the reason it is difficult to discuss these intermediate cases here derives from the coarse-grained categories we are using. In any case, I think that for the main goals of the paper, this coarse-grained analysis suffices.

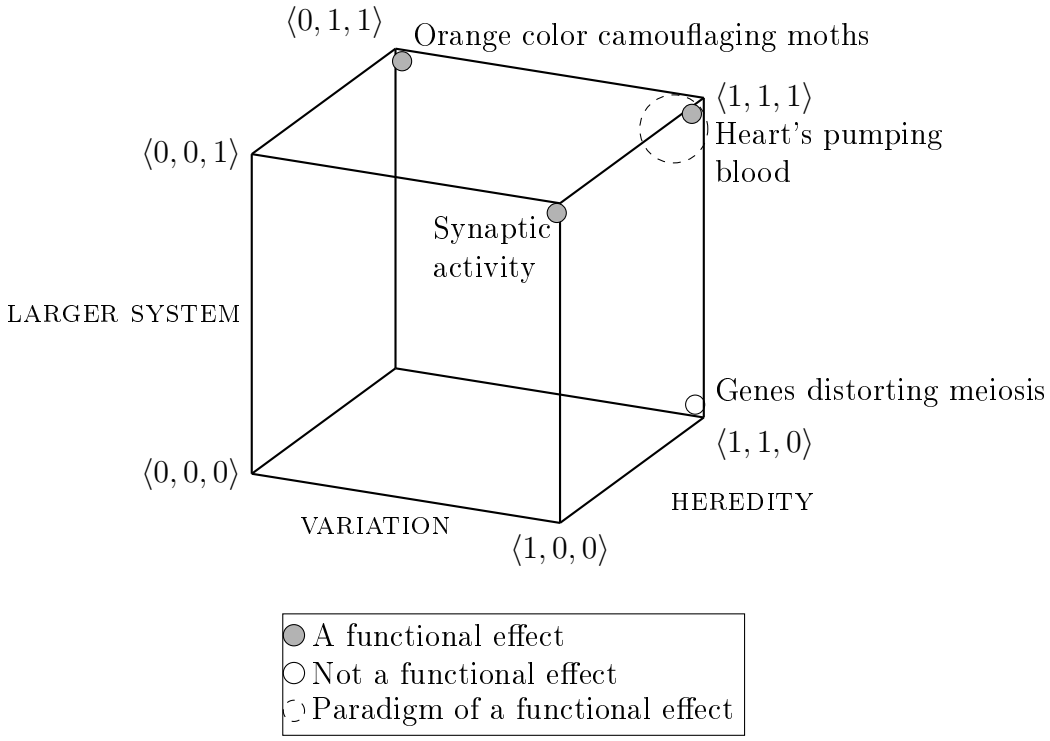

Figure 1: Effects classified according to heredity, variation and contribution to a larger system. Examples are discussed in the main text.

Wrapping up, in this section I argued that T3 can account for the properties associated with functions, once it is supplemented with the idea that functional effects must have contributed to the fitness of a larger system (i.e. T3+). Furthermore, I argued that functions are probably coextensional with effects satisfying T3+, and I also showed that this perspective can help to account for some alleged counterexamples. Thus, if T3+ is understood as describing the paradigm of function (from which marginal cases can differ in various respects), the arguments of this section lend support to T3+.

Recall, however, that our original goal was to assess the SEE Theory, which establishes a relationship between functions and the biological notion of selection for. Thus, in the next section we analyze the concept 'selection for' and in section 5 we will ascertain whether an analysis of this concept

goes hand in hand with the our best theory of functions. .

4 Selection For

In the previous section I argued for a certain interpretation of T3 (and T3+), namely that it should be understood as specifying the paradigm. To defend that interpretation, we began by identifying two properties associated with functions (their special explanatory and normative role), we described the connection between fulfilling the conditions stated in T3+ and possessing these two properties and we considered how some difficult cases can be classified as marginal instances of T3+. In this section, I propose to follow a parallel strategy with respect to T2. Why applying the same strategy? Recall that our goal is to assess whether the notion of biological function can be accounted for in terms of selection for; since I argued that T3+ should be understood as defining the paradigm (from which particular cases can differ in various respects), the obvious question is whether the analysis of selection for provided in T2 can also be understood in the same way. I will argue that it can, but (and here comes the crux of the argument), although paradigmatic instances of biological functions and selected effects largely overlap, this correspondence fails when we focus on marginal cases.

So let us apply the same strategy that we employed earlier. First of all, in the same way that described two important features of functions, i.e. Explanation and Normativity, we need to identify the explanatory role that the notion of selection is supposed to play.

So what are the explananda for the notion of *selection for*? What are the phenomena that this concept is supposed to account for? This issue remains controversial, but there seems to be broad agreement regarding two key ideas. On the one hand, *selection for* is a key concept in the theory of natural selection, and among other things natural selection provides persistence and origin explanations, i.e. natural selection explains the distribution of traits within a population and why some kinds of traits exist at all (see Neander, 1995, and Godfrey-Smith, 2014: 38).[19]. In this regard, the expanation of complex traits such as eyes or hearts plays a specially prominent role. On the other, *selection for* is a causal concept: Sober (1984) introduced it to distinguish those properties that play a causal role in the selection process

[19] A much more disputed question is whether natural selection can explain why particular individuals have certain traits.

from those that do not. The thick coat of polar bears, for instance, is selected for being warm, but a warm coat must be heavy, so there is also selection of heavy coats. In this case, there is selection *for* warmth, but only selection *of* thickness (Goode and Griffiths, 1995: 103).

Suppose for a moment that the main task of the notion of selection for is to identify a causal factor that accounts for the fact that a particular variant was favored over its alternatives. I think that all the conditions included in T2 are relevant for accommodating this property, but to keep things simple, I would like to briefly discuss two of them: *Variation* and *Heredity*.

First, notice that variation is necessary for selection, so variation in a population is essential for an effect F to be selected for. Furthermore, the origin of complex adaptations through selection requires changing the distribution of traits within a population, and this is only possible when individuals differ from each other (Neander, 1995). Therefore, in order to account for the origin and maintenance of certain types of traits, *selection for* clearly requires variation within a population.

In contrast with *Variation*, the idea that *Heredity* is necessary for natural selection (and, hence, for *selection for*) is much more controversial. Bouchard (2008, 2011), for example, has argued that evolution by natural selection does not require reproducing entities (so he challenges *Heredity* in T2). He points out that the evolution of fungi or plants, for instance, results in a change in somatic parts yet does not affect the germ line. One of the clearest examples he provides is the quaking aspen (*Populus tremuloides*), a tree that can reproduce both sexually and asexually. Asexual reproduction takes place when a root system shoots out what is called a 'runner' or 'sucker' that pierces the surface and grows as new stem. As a result, what looks like a normal forest with many individual trees may actually consist of one big clonal tree with interconnected branches. Bouchard argues that natural selection can take place in such a grove, e.g. a particular area can have a larger number of stems because it offers better ecological conditions. He also contends that this offers a causal explanation (Bouchard, 2008: 565).

I think there might be an important lesson to be learned from examples such as Bouchard's: in some cases, persistence might suffice for some kind of selection to take place.[20] An effect that does not clearly satisfy *Heredity*

[20]Bouchard defends a non-etiological theory of function called the 'Persistence Enhancing Propensity' Account (Dussault and Bouchard, 2017), but again, I think there is a way of accommodating these cases without departing from an etiological theory.

may nonetheless count as being selected for if it is an effect that causally accounts for differential persistence. This is not to deny, of course, that clear examples of complex adaptations fully satisfy *Heredity*; notice, for instance, that Bouchard's main example involves selection for location in the quaking aspen, which can hardly qualify as a complex adaptation. Thus, paradigm cases of *selection for* (e.g. the most explanatory) fulfill *Heredity*, but partial compliance with this condition might suffice for marginal instances.

Therefore, *Variation* is certainly required for a trait to be selected for, but a trait that does not satisfy *Heredity* might still qualify as a marginal case. Paradigm cases score highly on both of these axes. Needless to say, a full discussion of the explananda for a theory of natural selection, and the extent to which all the conditions in T2 are required, goes beyond the scope of this paper. Nonetheless, this will suffice for addressing our next and (final) question.

5 Selected-Effect Etiological Functions and Selection For

If the results of this paper so far are on the right track, then the relationship between having a biological function and being selected for becomes much more complex than usually thought. In particular, it follows that there is only a partial overlap between the process that grounds biological functions and the process that underpins selection. Let me elaborate.

To address the relationship between entities that evolve by natural selection and entities with functions, we can use the visual tool that we presented before, this time in a different way: instead of representing only functional entities, or only entities that are selected for, let us draw a three-dimensional space and represent both of them at the same time. In this way, we can visualize to what extent items that have a function coincide with items that are selected for, when certain dimensions are taken into account. Accordingly, let us consider three of the dimensions that we have argued are relevant for having a function: *Variation*, *Reproduction* and *Contribution to the fitness of a larger system*. Within this three-dimensional space, we will represent individuals that have a trait T with a function F as well as individuals that have a trait T that is selected for F. Following our previous conclusions, the result is depicted in figure 2.

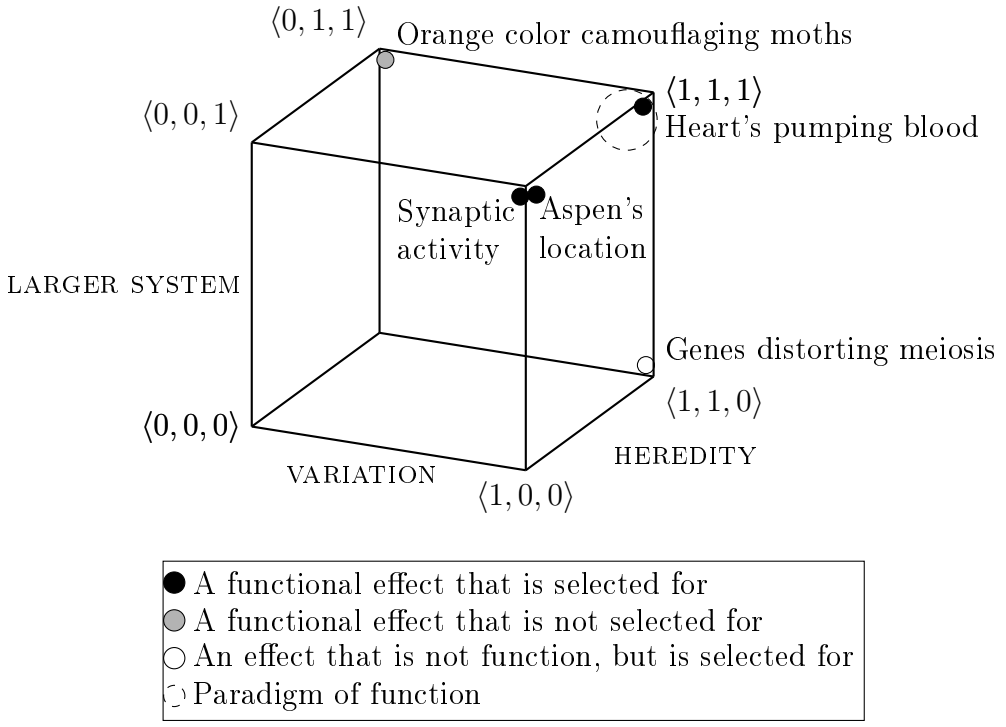

Figure 2: Entities with functions and selected effects, classified according to variation, heredity and contribution to the fitness of a larger system. Examples are discussed in the main text.

First of all, notice that paradigm instances of function probably overlap to a large extent with paradigm cases of selection for. Many items with a function F have actually been selected for F: hearts for pumping blood, wings for flying and eyelashes for protecting eyes (the dot representing hearts near the corner $\langle 1, 1, 1\rangle$ is supposed to indicate this). It is not hard to imagine that this considerable overlap between paradigm instances of function and selection has boosted the idea that function can be straightforwardly reduced to selection for.

Nonetheless, our previous discussion reveals that this match is not perfect: on the one hand, near the corner $\langle 1, 1, 0\rangle$ (i.e. variation and reproduction, but no contribution to a larger system) we will probably find some paradigm examples of natural selection, but not a single instance of function. The

reasons were presented earlier: there are no cases of function without contribution to the fitness of a larger system, but certainly these entities can still be selected for some effect F. Among other things, in this corner we would find segregation distortion genes or cancer cells, which fail to positively contribute to the fitness of any larger system but nonetheless evolve by natural selection.[21]

On the other hand, the presence of a grey dot near the corner $\langle 0, 1, 1 \rangle$ points to the existence of marginal cases of functional effects that have not been selected for. In particular, whereas variation is required for natural selection to occur, this might not be necessary for functionality (if Buller is right). Lewens' orange moths should be placed around this corner. Finally, some marginal cases of selection are also marginal instances of function, as the two dots at $\langle 1, 0, 1 \rangle$ illustrate. For instance, neural connections might get selected for through synapse competition, and also acquire functions without being reproductions of past neurons that had this specific effect. Similarly, quaking aspens might evolve by natural selection and possess functions, even if they clearly differ from the paradigm.

Two striking consequences seem to follow from this analysis. First of all, the set of entities that have a function and those that are selected for do not perfectly overlap. In other words, some items are selected for without having functions (e.g. cancer cells, segregation distortion genes), and others have functions without being selected for (e.g. orange moths). Relying on some ideas discussed in this paper, one way of explaining this mismatch appeals to the different explanatory goals of a theory of natural selection and the theory of function: among other things, the former is supposed to explain the distribution of traits and the origin of complex adaptations, whereas the latter is supposed to account for Normativity and Explanation. This partial overlap might be a consequence of the divergent explanatory aims.

At the same time, however, is also remarkable that paradigm instances of both kinds of entities probably occupy a similar region of space. I do not think that this is an artificial result deriving from the three dimensions considered in this paper; as I suggested above, this significant correspondence can explain why the idea of reducing function to selection for has been so attractive. Both selection for and function seem to be grounded in similar

[21] This is not to say that cancer cells are paradigm cases of function *all things considered* (see Lean and Plutynski, 2016; Cf. Germain, 2012, Germain and Laplane, 2017). Figure 2 highlights the area that paradigm cases of function occupy only with respect to three dimensions: variation, heredity and contribution to the fitness of a larger system.

processes. Nonetheless, the results of this essay suggest that the relationship between them is more complex than has usually been supposed.

This outcome raises a final question, whose answer might have important consequences not only for philosophy, but also for many projects in cognitive science, biology and medicine: does this result imply that the notion of function cannot be accounted for in terms of natural selection? Does it undermine etiological and naturalistic theories of function? Certainly, the conclusions of this paper challenge the idea that biological functions can simply be reduced to effects being selected for; that is, they question T1. Hence, any simplistic approach that merely seeks to identify functions with selected effects is probably doomed to failure. Nonetheless, I do not think the naturalization of functions is in jeopardy.

There are at least two ways in which the strategy pursued by etiological theories of function can be vindicated. On the one hand, even though our previous results show that functional effects cannot be straightforwardly reduced to selected effects, the project of reducing functions to some causal process that took place in the recent past is probably on the right track. Indeed, it is very plausible that some process approximating natural selection grounds function attributions: as I argued in section 3, this is the process that is most likely to be able to ground NORMATIVITY and EXPLANATION, among other things. Consequently, an etiological theory slightly different from the SEE approach can be vindicated. I argued this account of function will probably include contributing to the fitness of a larger system as a requisit, and it will not require variation in the population (although variation might be still be needed for qualifying as a paradigm case). Thus, it will count as a form of etiological theory (e.g. functions are defined as past effects that contributed to a larger system in certain ways), but probably not as a version of the selected-effects approach, as it is usually understood.

Secondly, notice that the set of paradigm entities that have a function significantly overlaps with the set of paradigm entities that are selected for. This is not a minor point, since many generalizations would still hold on the basis of this relationship. Indeed, for many purposes, identifying functional entities with items that have been selected for might still be good enough. That could be an admissible simplification, as long as we keep in mind that the real relationship is probably more complex.

6 Conclusions

Time to sum up. I formulated the Selected-Effect Etiological Theory of function as involving two claims, T1 and T2, and a logical consequence, T3. The results of this paper show that we need to revise all three of them.

First, I argued that T3 should probably be supplemented with a fourth condition, namely that a trait should contribute to the fitness of a larger system (what I called 'T3+'). The main motivation is to provide a theory that be able to account for the two main features of functions (EXPLANATION and NORMATIVITY) and address some counterexamples. Secondly, I argued that T2 and T3+ should primarily be understood as establishing what is required for an effect to be a paradigm case of selection for and function, respectively. Hence, although prototypical instances of function and selection for must satisfy all of these conditions, marginal cases might fail to comply with some of them. Thirdly, some of the conditions in T2 and T3+ are necessary conditions for being an instance of function or selection for. For example, no entity that fails to comply with *Contribution to a Larger System* can be said to have a function (not even marginally) and no effect that fails to fulfill *Variation* can be said to have been selected for. I also argued that this perspective can accommodate some alleged counterexamples that have been raised against the standard analyses of these two notions. Additionally, I contended that on this interpretation, the process that gives rise to functions only partly overlaps with the selection process involved in natural selection. In other words, some of the dimensions that are relevant for attributing one property are also relevant for attributing the other, but not all of them. Thus, T1 should be abandoned. This might be partly due to the different explanatory goals of a theory of natural selection and a theory of function.

Finally, I suggested that this perspective is fully compatible with a naturalistic theory of function, since it only undermines the simplistic view that seeks to reduce one property to the other. The door is still open for an approach that explains functions in terms of a causal process from the recent past that approximates natural selection in various respects. Thus, the results of this paper vindicate an etiological account of function, albeit one that must differ from the Selected-Effects Etiological Theory.

Acknowledgments

I would like to thank Ruth Millikan, Fermín C. Fulda, Justin Garson, Javier Gonz ález de Prado Salas, John Horden, Fabian Hundertmark, Ulrich Krohs, Brian Leahy, Christian Nimtz, Cailin O'Connor, Peter Schulte, Ema Sullivan-Bisset, Hannah Rubin and an anonymous reviewer for their helpful comments and criticisms. Earlier versions of this paper were presented at the workshop 'Teleosemantics and the Nature of Functions' at the University of Bielefeld in September 2017 and the conference 'The Generalized Theory of Evolution' at the University of Duesseldorf in February 2018. Financial support was provided by the MINECO project 'la Complejidad de la Percepción: Un Enfoque Multidimensional' (FFI2014-51811-P) and 'Varieties of Information' (PGC2018-101425-B-I00).

Conflict of Interest

The author declares that they have no conflict of interest.

References

[1] Abrams, M. (2005) Teleosemantics Without Natural Selection. *Biology and Philosophy.* 20: 97-116

[2] Artiga, M. (submitted) A Dual-Aspectr Theory of Artifact Functions.

[3] Bedau, M. (1991) Can Biological Teleology Be Naturalized?. *Journal of Philosophy.* 88(11):647-655.

[4] Birch, J. (2012b) Robust Processes and Teleological Language. *European Journal for Philosophy of Science.* 3:299-312.

[5] Boorse, Ch. (1976) Wright on Functions. *The Philosophical Review*, 85(1):70-86.

[6] Boorse, Ch. (2002) A Rebuttal on Functions. In A. Ariew and M. Perlman, *Functions*, OUP.

[7] Bouchard, F. (2008) Causal Processes, Fitness, and the Differential Persistence of Lineages. *Philosophy of Science,* 75(5): 560-570.

[8] Bouchard, F. (2011) Darwinism Without Populations: A More Inclusive Understanding of the 'Survival of the Fittest'. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences,* 42(1):106-114.

[9] Buller, D. (1998) Etiological Theories of Function: A Geographical Survey. *Biology and Philosophy,* 13: 505-527.

[10] Buller, D. (2002) Function and Design Revisited. In A. Ariew, R. Cummins and M. Perlman, (eds.) *Functions,* OUP.

[11] Canfield, J. (1963) Teleological Explanation in Biology. *British Journal for the Philosophy of Science,* 14(56): 285-295

[12] Cummins, R. (1975) Functional Analysis. *Journal of Philosophy,* 72: 741-765

[13] Cummins, R. (2002) Neo-teleology. In Andre Ariew, Robert Cummins and Mark Perlman (eds.) *Functions: New Essays in the Philosophy of Psychology and Biology*, OUP.

[14] Dussault, A. and F. Bouchard (2017) A persistence enhancing propensity account of ecological function to explain ecosystem evolution. *Synthese*, 194: 1115–1145.

[15] Garson, J. (2011) Selected effects and causal role functions in the brain: the case for an etiological approach to neuroscience. *Biology and Philosophy,* 26(4): 547-565.

[16] Garson, J. (2012) Function, selection, and construction in the brain. *Synthese,* 189(3): 451-481.

[17] Garson, J. (2016) *A Critical Overview of Biological Functions,* Springer.

[18] Garson, J. (2017) A Generalized Selected Effects Theory of Function. *Philosophy of Science,* 84(3): 523-543.

[19] Garson, J. and D. Papineau (2019) Teleosemantics, Selection and Novel Contents. *Biology and Philosophy.* 34: 36.

[20] Garson, J. (forthcoming) There Are No Ahistorical Theories of Function. *Philosophy of Science.*

[21] Germain, P.(2012) The Evolution of Failure: Explaining Cancer as an Evolutionary Process. *Biology and Philosophy,* 27 (6):785-810.

[22] Germain, P. and L. Laplane (2017) Metastasis as supra-cellular selection? A reply to Lean and Plutynski. *Biology and Philosophy,* 32 (2):281-287.

[23] Goode, R. and P. Griffiths (1995) The Misuse of Sober's Selection for/Selection of Distinction. *Biology and Philosophy,* 10 (1): 99-108.

[24] Griffiths, P. (1993) Functional analysis and proper functions. *British Journal for the Philosophy of Science,* 44(3): 409-422.

[25] Godfrey-Smith, P. (1993) Functions: consensus without unity. *Pacific Philosophical Quarterly,* 74(3): 196-208.

[26] Godfrey-Smith, P. (1994) A Modern History Theory of Functions. *Noûs,* 28(3): 344-362

[27] Godfrey-Smith, P. (2009) *Darwinian Populations and Natural Selection,* OUP

[28] Godfrey-Smith, P. (2014) *Philosophy of Biology,* Princeton University Press

[29] Huber, J. D. and R. Inglehart (1995) Expert Interpretations of Party Space and Party Locations in 42 Societies. *Party Politics* 1 (1): 73–111.

[30] Jackon, F. and P. Pettit (1990) Program explanation: a general perspective. *Analysis* 50 (2): 107–117.

[31] Krange, O., B. P. Kaltenborn and M. Hultman (forthcoming) Cool dudes in Norway: climate change denial among conservative Norwegian men. *Environmental Sociology.*

[32] Kreamer, D. (2014) Revisiting recent etiological theories of functions. *Biology and Philosophy.* 29(5): 747-759

[33] Lean, Ch. and A. Plutynski (2016) The Evolution of Failure: Explaining Cancer as an Evolutionary Process. *Biology and Philosophy* 31 (1):39-57.

[34] Lewens, T. (2004) *Organisms and Artifacts Design in Nature and Elsewhere.* MIT Press

[35] Lewontin, R. (1985) Adaptation. In R. Levins and R. Lewontin(eds.) *The Dialectical Biologist,* Harvard University Press, 65-84.

[36] Maley, P. and G. Piccinini (2017) A Unified Mechanistic Account of Teleological Functions for Psychology and Neuroscience. In David Kaplan (ed.), *Explanation and Integration in Mind and Brain Science*, OUP: 236-256.

[37] Manning, R. (1997) Biological Function, Selection, and Reduction. *British Journal for the Philosophy of Science ,* 48(1): 69-82

[38] McCright, A. M. and Dunlap, R. E. (2011) Cool dudes: The denial of climate change among conservative White males. *Global Environmental Change*, 21, 1163-1172

[39] Millikan, R. (1984) *Language, Thought and Other Biological Categories.* MIT Press

[40] Millikan, R. (1989) In Defense of Proper Functions. *Philosophy of Science,* 56(2): 288-302

[41] Millikan, R. (1993) *White Queen Pscyhology and Other Essays for Alice.* MIT Press

[42] Millikan, R. (2002) Biofunctions: Two Paradigms. In A. Ariew and M. Perlman (eds.) *Functions.* OUP.

[43] Millikan, R. (2017) *Beyond Concepts.* OUP

[44] Mossio, M., C. Saborido and A. Moreno (2009) An Organizational Account of Biological Functions. *British Journal for the Philosophy of Science,* 60(4): 813-841

[45] Neander, K. (1991a) Functions as selected effects: The conceptual analyst's defense. *Philosophy of Science,* 58(2): 168-184

[46] Neander, K. (1991b) The Teleological Notion of 'Function'. *Australasian Journal of Philosophy,* 69(4): 454-468

[47] Neander, K. (1995) Pruning the Tree of Life. *British Journal for the Philosophy of Science,* 46 (1):59-80.

[48] Neander, K. (2017) Functional analysis and the species design. *Synthese,* 194:1147-1158.

[49] Neumayer, E. (2004) The Environment, Left-Wing Political Orientation And Ecological Economics. *Ecological Economics* 51 (2004): 167-175

[50] Papineau, D. (2001) The status of teleosemantics, or how to stop worrying about swampman. *Australasian Journal of Philosophy,* 79 (2): 279-89

[51] Price, C. (1995) Functional Explanations and Natural Norms. *Ratio,* 8(2): 143-160

[52] Price, C. (2001) *Functions in Mind: A Theory of Intentional Content.* OUP

[53] Queller, D. and J. Strassman (2009) Beyond society: the evolution of organismality *Philosophical Transactions of the Royal Society.* 364: 3143-3155

[54] Riddley, M. (2004) *Evolution.* Blackwell Publishing

[55] Schulte, P. (forthcoming) Why Mental Content is Not Like Water: Reconsidering the Reductive Claims of Teleosemantics. *Synthese*

[56] Schwartz, P. (1999) Proper Function and Recent Selection. *Philosophy of Science,* 66: S210-S222

[57] Schwartz, P. (2002) The Continuing Usefulness Account of Proper Functions. In A. Ariew and M. Perlman (eds.) *Functions*, OUP.

[58] Shea, N. (2018) *Representation in Cognitive Science.* OUP

[59] Sober, E. (1984) *The Nature of Selection.* University of Chicago Press

[60] Sober, E. (1995) Natural selection and distributive explanation: A reply to Neander. *British Journal for the Philosophy of Science,* 46 (3):384-397

[61] Sorabji, R. (1964) Function. *The Philosophical Quarterly,* 14(57): 289-302

[62] Sterelny, K. (1995) Basic Minds. *Philosophical Perspectives,* 9: 251-270

[63] Woodfield, A. (1976) *Teleology.* Cambridge University Press

[64] Wouters, U. (2005) The Function Debate in Philosophy. *Acta Biotheoretica,* 53(2):123-151

[65] Wright, L. (1973) Functions. *The Philosophical Review,* 82(2):139-168