Do Proper functions explain the existence of traits?

Abstract

The theory of Proper functions has been developed with a simple and intuitive view of natural selection. Traits are selected because of their positive effects on the fitness of the organisms that have them. These ‘selected effects’ are the Proper functions of the traits. This thought underpins the key attraction of the theory, namely that the Proper function of a trait is also a causal explanation of the trait: the trait exists because it performs this function. We show, however, that Proper functions as currently defined explain the existence of traits only under highly restrictive assumptions about evolutionary dynamics. In many common scenarios in which traits evolve by natural selection, Proper functions do not explain those traits. This is because definitions of Proper Function extract from any evolutionary scenario only the information that would be explanatorily relevant in the simplest evolutionary scenario. In more complex scenarios Proper functions omit by design the information that is explanatorily relevant in those scenarios. The assumptions required for Proper functions to be explanatory are particularly unlikely to hold in the domain that Proper function theorists care most about - the evolution of mind and behavior. A more sophisticated theory of Proper functions may be possible, but recent philosophy of population biology has identified serious obstacles that such a project would have to overcome.

Introduction

Teleological language in biology, such as the ‘function’ or ‘purpose’ of a biological character, or the ‘goal’ of an activity, can be interpreted as an elliptical reference to natural selection. In the neo-Darwinian modern synthesis this way of thinking about function, purpose and goal-directedness was widely accepted and was known as ‘teleonomy’ (Pittendrigh 1958). The idea was nicely expressed by the founder of ethology, Konrad Lorenz:

"If we ask 'What does a cat have sharp, curved claws for?' and answer simply 'To catch mice with', this does not imply a profession of any mythical teleology, but the plain statement that catching mice is the function whose survival value, by the process of natural selection, has bred cats with this particular form of claw."(Lorenz 1966, p. 9)

In philosophy, this idea underpins the theory of ‘Proper functions’, also known as the ‘etiological’ or ‘selected effects’ theory of functions, which originated in Ruth Millikan’s influential Language, Thought and Other Biological Categories (Millikan 1984), and in Karen Neander’s influential but unpublished PhD thesis and subsequent papers (Neander 1983, 1991a, b). “The gist of this…
theory is that the (or a) function of an item... is what it was selected to do.” (Neander 2017, p. 20)

Advocates of the theory of Proper functions have emphasised that what distinguishes it from other interpretations of teleological language in biology is that Proper functions causally explain why traits exist. The brilliant colours of the male Superb Fairy Wren exist today because ancestral wrens with brilliant colours survived and reproduced more effectively than they would have without them. In ancestral wrens, the colours increased fitness by attracting mates, so in today’s wrens the Proper function of these colours is to attract mates. The wren’s colours have other effects and some of these may benefit the wren or benefit some other organism, but only of this specific effect can we say that this is what the colour is for; that the colour exists because it has this effect. This special explanatory role is what makes it the ‘Proper’ function of the trait.

Neander states very clearly that this explanatory role is the basic rationale for the selected effects theory of Proper function: "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, p. 457). She continues, "...my view is that function attributions universally and intrinsically justify teleological explanations" (Neander 1991b, p. 458). They do so as a result of, “a causally explanatory selection process, during which those items or traits were selected for those effects which are their functions” (Neander 1991b, p. 467).

Leaping forward twenty-eight years, Justin Garson, the current leading recent advocate of the theory, makes the fact that functions causally explain traits his master argument in favour of the theory of Proper functions:

“consider a puzzling feature of ordinary biological usage, namely, function’s explanatory depth: sometimes function statements are causal explanations for traits. When biologists say the function of the zebra stripes is to deter biting flies, for example, they’re trying to explain why zebras have stripes. … If we take explanatory depth seriously, then the traditional selected effects theory, or something in its neighborhood, has no equal.” (Garson 2019, p. 7)

Our concern in this paper is that Proper functions as currently defined only explain the existence of the traits that bear them under restrictive assumptions about evolutionary dynamics. Proper functions explain traits only in the absence of feedback from natural selection to the relative fitness of the trait. Such feedback generates environmental heterogeneity that is endogenous to the selection process. But even simple, exogenous environmental heterogeneity can also lead to scenarios in which Proper functions do not explain traits. Both endogenous and exogenous environmental heterogeneity are ubiquitous in evolution. Proper function theorists have not had to confront these problems because they use simple, intuitive examples of natural selection presented via verbal scenarios. Typically, these scenarios assume a uniform selective

---

1 An extensive philosophical literature compares this interpretation of functional language in biology to alternative interpretations. For a recent review see (Garson 2016).
environment which precludes trait-environment feedback. We argue that this is neither a useful abstraction nor a reasonable simplifying assumption. Instead it obscures important evolutionary dynamics that undermine the generality of the theory of Proper functions. Under many realistic evolutionary scenarios, attempting to explain a trait by saying that it is there because of its Proper function omits the key factor or factors in the causal explanation of the existence of the trait.

The simple, intuitive picture of natural selection on which selected effects theorists have relied is one in which traits evolve because they solve some specific problem or problems. It is the Proper function of the trait to solve those problem(s) and they exist because they solve them. So the Proper function of the trait explains why it exists. Take the oft-used example of “the function of the heart is to pump blood”. Selected effects theorists take it that there once was an organism with no heart (or with a poor heart) which gave rise to an organism with a heart (or with a better heart). On account of its ability to pump blood, thus distributing nutrients and removing waste products more efficiently than its predecessor, the new heart conferred a relative advantage on its possessor. Selected effects theorists assume that interactions between individuals -- organisms with hearts outcompeting organisms with no hearts or poorer hearts -- aggregate additively and cause organisms with hearts to increase in frequency and eventually fix in the population. But the trajectories of populations under natural selection only conform to this picture under the restrictive assumptions that traits have no influence on the environment. When traits and environments interact, traits evolve in the face of an ever changing series of problems, where the nature of the problem at any one point in time is a function of the evolution of the trait to that point. As we will show, evolution is often like this. In other cases, environments are exogenously heterogenous, and the evolution of traits depends not just on their relative performance in patches where they do well, but on their relative performance in patches where they do badly. So it is not generally true that the adaptive trajectory of a trait in a population ('why it exists') is explained by the fact that this trait has repeatedly solved a specific problem (or set of problems) better than the alternatives.

In the next two sections we examine current definitions of Proper function in detail. In sections four and five we substantiate our criticism of the theory by demonstrating it in some realistic evolutionary scenarios.

**Selected effects accounts of Proper function**

There are many definitions of ‘Proper function’ in the philosophical literature. Some influential examples include:

(Defn. 1.) “...for an item A to have a function F as a "proper function", it is necessary (and close to sufficient) that... A originated as a "reproduction" (to give one 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 this or these performances. (Millikan 1989, p. 288)
(Defn. 2.) “It is a/the 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.” (Neander 1991a, p. 74).

(Defn. 3.) “Where i is a trait of systems of type S, a proper function of i in S's is F iff a proximal selective explanation of the current non-zero proportion of S's with i must cite F as a component in the fitness conferred by i.” ['proximal' means recent episodes of selection] (Griffiths 1993, p. 415).

(Defn. 4.) “The function of m is to F iff:
(i) m is a member of family T,
(ii) members of family T are components of biologically real systems of type S,
(iii) among the properties copied between members of T is property or property cluster C, which can do F,
(iv) one reason members of T such as m exist now is the fact that past members of T were successful under selection in the recent past, through positively contributing to the fitness of S, and
(v) members of T were selected because they did F, through having C.”
(Godfrey-Smith 1994, p. 359)

(Defn. 5) Shea defines Proper function as part of a longer definition of when a trait (R) has the Proper function of carrying information about a state of affairs (C). The relevant clause of Shea’s definition requires that, “an evolutionary explanation of the current existence of the representing system adverts to Rs having carried information about C” (Shea 2007, p. 419)

The differences between these definitions are dwarfed by their similarities. All share the idea that a Proper function is an effect whose positive effect(s) on fitness plays a role in explaining why a trait evolved by natural selection. Exactly what role the positive effect on fitness must play in an evolutionary explanation is less clear from these definitions. Shea only says that the evolutionary explanation ‘adverts to’ the effect. Neander says that the effect “caused the genotype... to be selected by natural selection” (Neander 1991a, p. 174). Godfrey-Smith says that the trait was “successful under selection in the recent past, through positively contributing to... fitness” (Godfrey-Smith 1994, p. 359). Both Neander and Godfrey-Smith’s formulations are naturally interpreted as requiring that the effect caused an increase in the proportion of the trait relative to its alternatives. Griffiths takes a slightly different tack – a selective explanation of the observed proportion of the trait needs to include this effect on fitness if it is to succeed in explaining the observed proportion. But all the definitions agree that the positive effects of the trait on individual fitness can be used to explain the representation of the trait in the target population. We will show that the vagueness of these claims about natural selection have obscured a major problem with selected effects theory: in many realistic and important evolutionary scenarios, Proper functions do not explain the traits that bear them.
The most careful definition of Proper function in the philosophical literature is undoubtedly that of Millikan (1984). Definitions 1-5 all state or imply that the trait which has the proper function is the same trait, in some sense, as the trait that was selected in ancestral populations. Millikan spells this out in more detail with her definition of a ‘reproductively established family’ of trait tokens. What Millikan calls a ‘family’ of traits biologists call a ‘homologue’ - a set of traits related to one another by descent from a common ancestor.

Having defined ‘reproductively established family’, Millikan defines proper function as follows:

(Defn 6) “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 directly causing reproduction of m or explaining why R was proliferated and hence why m exists.” (Millikan 1984, p. 28)

Speaking from experience, we can say that the first response of evolutionary biologists to this definition is that the terminology is unfamiliar and confusing. A table of equivalences between Millikan’s terminology and the conventional terms in biology is given as an Appendix. Restating Millikan’s definition using conventional biological terms will make it easier to apply her theory to realistic biological examples. These terms are organisms (we could say “genes, organisms or

---

2 “First we define the notion ‘ancestor of a member of a reproductively established family’:
(1) Any member of a (first-order) reproductively established family from which a current member m was derived by reproduction or by successive reproductions is an ancestor of m.
(2) Any temporally earlier member of a (higher-order) reproductively established family which member was produced by an ancestor of the device that produced a present member m is an ancestor of m.
(3) Any earlier member of a (higher-order) reproductively established family that a present member m is similar to in accordance with a proper function of a producer that produced both is an ancestor of m.”
(Millikan 1984, pp. 27–28)

3 Millikan’s ‘reproductively established families’ are not lineages of organisms (e.g. wrens) or even lineages of characters in the familiar sense of a ‘transformation sequence’ of characters such as the sequence from scales to feathers. Instead, they are lineages of individual characters: one blue head is the ‘ancestor’ of another blue head and all the homologous blue heads are a ‘family’.

4 Millikan’s distinction between ‘direct’ and ‘derived’ functions will not concern us here. Very briefly, her idea is that objects which do not have a history of natural selection, such as the stones in the case of a Caddis fly larvae, can nevertheless have Proper functions ‘derived’ from the Proper function of the mechanisms that make the case.
groups”), populations and characters. In comparative biology ‘character’ refers to what philosophers call a ‘determinable property’ and ‘character state’ to what philosophers call a ‘determinate property’. For example, head colour is a character of which blue and brown are possible character states. Turning blue for the breeding season is a character of which March to October and Never are two possible character states.

Our proposed canonical statement of the selected effects theory of Proper function restates Millikan’s definition using standard terminology in evolutionary biology:

(Defn. 7.) Consider a population of organisms some of whom have character state C_i. These organisms are descendants of organisms with C_i and their character states are homologous. In this population, C_i has effect F as a direct Proper function if and only if:

7.1. In some ancestral populations there was variation in C
7.2. Having state C_i caused some ancestral individuals to produce effect F with higher probability than individuals with alternative character states
7.3. Performing F caused some ancestral individuals with C_i to have greater reproductive output than they would have had if C_i had been changed to one of the extant alternative character states in that ancestral population (C_i-C_n)
7.4. The frequency of C_i in the current population is explained by selection for C_i in these ancestral populations in virtue of the increase in fitness accruing to individuals with C_i from performing F more frequently than individuals with alternative character states

From Defn. 7, we can clearly see the general structure of the selected effects account:

- Condition 7.1 is necessary if natural selection is to operate.
- Condition 7.2 establishes a causal connection between our focal trait and its function.
- Condition 7.3 mechanistically connects our focal trait and its function with a fitness advantage over individuals with alternative character states.
- Condition 7.4 connects our trait, its function, and its fitness advantage over competing traits, with the spread of the trait in the population.

When all four conditions are satisfied, F is the Proper function of the trait.

Conditions 1-3 in Defn. 7 are all deliberately weak. There is some variation which correlates to some degree with the effect and this has some effect on reproduction. Condition 4 can be read either strongly or weakly, and this is where Proper function theorists go wrong. The weak reading is merely that the facts established in 1-3 have some explanatory relevance to the evolution of the trait. If these facts were different, that would have some effect on the evolution of the trait. As we will see, this is usually if not always true. The strong reading is that we can explain the trait by citing its Proper function: the facts established in 1-3 are not merely some explanatorily relevant facts, they are the key explanatory facts. Any other explanatorily relevant facts can accordingly be regarded as background conditions, or included in a ceteris paribus
clause. For example, it might be necessary that the effect identified in 1-3 happens often enough, or that there be no other, countervailing fitness reducing effects of the trait. The fact that a Proper function explanation was incomplete in that sort of way would not be a serious worry. But in the cases we review below, identifying the Proper function of a trait is never more than a small step towards an evolutionary explanation of why the trait exists. Key explanatory facts of equal or greater centrality are simply omitted from the explanation that cites the Proper function. This undermines the key attraction of the theory of Proper functions, the idea that the Proper function(s) of a trait explains the prevalence of that trait.5

3. When do Proper functions explain traits?

3.1 Zebra stripes and biting flies

We will begin by exploring Garson’s example: “the function of zebra stripes is to deter biting flies” (Garson 2019). We start by making clear what is assumed here. First, Garson assumes that biting flies are less likely to bite animals with stripes. There is evidence for this both from experiments (Caro et al. 2019) and phylogenetic correlations (Caro et al. 2014). Second, Garson assumes that the selective force for zebra stripes comes about through an interaction between zebras and biting flies. Beyond that, we claim, little is known or assumed by Garson or other philosophers of biology who use this example. They do not specify the evolutionary dynamics involved in the spread of the stripe trait, and in particular how the dynamics might be influenced by coevolution with biting flies. We will show how these unspecified elements affect whether or not Proper functions are also causal explanations.

As per Defn 7, let the character C be Pattern, of which we assume there are only two variants: the striped trait C_1 (denoted Stripes), and the non-striped trait C_2 (denoted Non-stripes). Furthermore, we assume that the effect F of Stripes is to reduce the probability of being bitten by biting flies (denoted Reduce_Bites), which it does more efficiently than Non-stripes. We can now fill out Defn. 7 as follows.

Consider a population of organisms some of whom have Stripes. These organisms are descendants of organisms with Stripes and their character states are homologous. In this population, Stripes has effect Reduce_Bites as a direct Proper function if and only if:

1. In some ancestral populations there was variation in Pattern
2. Having Stripes caused some ancestral Striped individuals to Reduce_Bites with higher probability than Non-striped individuals.

---

5 Selected effects proponents could respond that when condition 7.4 is only weakly met, traits do not have Proper functions. This would save the idea that Proper functions really explain traits at the expense of limiting the application of the theory. In particular, the idea would have little if any application to the main interest of Proper function theorists: mind, behavior and communication (Section 5).
3. Reduce_Bites caused some ancestral Striped individuals to have greater reproductive output than they would have had if Stripes had been changed to Non-stripes in that ancestral population.

4. The frequency of Stripes in the current population is explained by selection for Stripes in these ancestral populations in virtue of the increase in fitness accruing to Striped individuals from Reducing_Bites more frequently than Non-striped individuals.

So far so good. We have causally connected Stripes with Reduce_Bites (7.2), Stripes and Reduce_Bites with a fitness advantage over Non-stripes (7.3), and the fitness advantage of Stripes and Reduce_Bites with Stripes replacing Non-stripes in the population (7.4).

We believe that the account above is a faithful explication of what selected effects theorists mean when they say “the function of zebra stripes is to deter biting flies” (Garson 2019). Given that Defn 7 is fully specified, this is the only selected effect explanation that we can apply in this case. Nevertheless, details of the coevolutionary dynamics between zebra stripes and biting flies have been left unspecified and do not feature in the selected effects explanation. We first show a set of assumptions under which the Proper function explanation accurately reflects the underlying evolutionary explanation. Later, we show how coevolutionary interactions would change the underlying evolutionary explanation, causing the Proper function explanation to no longer reflect the evolutionary explanation for zebra stripes.

3.1.1 No co-evolutionary dynamics: Proper function gets it right

Let us assume a single homogeneous environment in which biting flies are always present and in which their bites are the only factor acting on zebra reproduction and mortality. Biting flies have a predilection for Non-striped individuals and this preference does not evolve. (We might imagine that biting flies have developmental or genetic constraints that preclude them from being able to co-evolve with respect to the Pattern of their hosts.) We can suppose in an ancestral population of Non-striped individuals, there once arose a Striped individual through mutation, who on account of being subject to fewer biting fly bites, was able to survive and reproduce more efficiently than Non-striped individuals, a process that continued unabated until all Non-striped individuals had been replaced by Striped individuals. In this case a selected effects explanation gets it right, accurately reflecting the underlying evolutionary explanation, because the selective regime is invariant with respect to the advantage of Striped individuals over Non-striped individuals.

3.1.2 Co-evolutionary dynamics: Proper function gets it wrong

Let us now assume that biting flies can evolve their preferences for Pattern (assume that a single fly either specialises on Stripes or on Non-stripes). We assume that biting flies can evolve rapidly and that the more frequent a Pattern, the stronger the selective force acting on biting
flies to specialise upon it⁶. Let us assume that when Stripes evolved, the majority of mammal species upon which the biting fly fed were Non-striped (which has support in modern populations; (Jordan et al. 1962). We must now consider the possibility that Stripes causes Reduce_Bites not because of inherent constraints on biting fly evolvability but rather because of the historical context of biting flies. Biting flies happened to evolve to specialise on Non-striped species and zebras (and any other Striped species) happened upon a rare phenotype to which the biting fly was not coadapted. In other words, Stripes Reduce_Bites because it is a rare phenotype. It follows that if Stripes were to rise in frequency and become the dominant Pattern, the biting fly would respond by evolving to specialise on Stripes. Once this happens, Stripes would increase bites, creating problems for our causal explanation of the existence of stripes.

Such coevolutionary dynamics between hosts and parasites are well-studied in the context of Red Queen dynamics (Brockhurst et al. 2014). To simplify the comparison with the selected effects explanation, we will describe the Red Queen dynamics assuming that zebras are the only host of biting flies. If the parasite and host evolve at a similar rate—such that host-parasite coevolutionary feedbacks occur over a short timescale—we expect to see fluctuating selection (the frequency of Stripes will oscillate from low to high along with a time-lagged oscillation of the frequency of biting fly Stripes preference) (Brockhurst et al. 2014). Given these altered dynamics, the relationship between the clauses of Defn. 7 now changes as follows.

We have causally connected Stripes with Reduce_Bites contingent upon Stripes (and thus the biting fly’s preference for Stripes) being rare (7.2), Stripes and Reduce_Bites with a fitness advantage over Non-stripes so long as Stripes (and thus the biting fly’s preference for Stripes) remain rare (7.3), and the frequency-dependent fitness effect of Stripes and Reduce_Bites with Stripes coexisting in a fluctuating equilibrium with Non-stripes in the population (7.4).

This new description contains a number of conditional statements and has lost the clean association between Stripes, Reduce_Bites, and a causal story of natural selection. To claim that the Proper function of Stripes is to Reduce_Bites is clearly problematic when roughly half the time Stripes do not Reduce_Bites but actually increase them. There is, however, no easy way out of this quandary because our selected effects formulation is fully specified and cannot be altered to account for coevolutionary dynamics.

The point of this discussion is not to argue about how zebra stripes actually evolved! We simply want to highlight how implicit assumptions about evolutionary dynamics are made whenever a Proper function is cited as explanation. Whether we actually need to consider coevolutionary dynamics in an evolutionary explanation of zebra stripes is not important. Numerous examples of natural selection to which we might hope to apply a Proper function explanation do involve coevolutionary dynamics (e.g. parasitic cuckoos (Servadio and Lande 2003), plant-insect mutualisms (Bronstein et al. 2006), and as we will discuss in section 4.1, the evolution of

---

⁶ Selection on biting fly evolution increases with the frequency of a Pattern because the more common a Pattern, the more frequent the encounter rate between flies and that type, which increases the advantage an individual fly would gain by specialising on that Pattern.
signalling and communication. If the theory of Proper functions is to live up to its promises, then it must be able to accommodate these sorts of evolutionary dynamics. If not, its proponents must admit that it is a theory that can at best apply to a subset of traits, while failing wherever evolutionary feedback occurs. In the next subsection, we show how negative frequency-dependence precludes Proper function explanation in a realistic biological example.

3.2 Frequency-dependent selection in the Gouldian finch

In this section we show in more detail using a real case how frequency dependence undermines the idea that traits exist because of their Proper functions. We start with the simple, textbook evolutionary scenario known as the ‘Hawk-Dove game’.

In this game the two types can persist in the population despite the fact that every time the two types interact the payoff to Hawks exceeds the payoff to Doves. Individual Hawks always beat individual Doves, but the costs to Hawks of their interactions with other Hawks are so high that as the proportion of Hawks in the population rises the average payoff to H comes down and the average payoff to D goes up. The equilibrium ratio of H to D is such that the average fitness of the two types is equal.

Fig 1. Payoff matrix for the Hawk-Dove game.

A natural example of this evolutionary game has been observed in the spectacular Gouldian Finch of northern Australia (Kokko et al. 2014). Simplifying somewhat, the species has two morphs, red-headed and black-headed. The two morphs mate assortatively, their mating preference sustained by the fact that hybrid offspring have reduced fitness. Red-heads have elevated production of testosterone and corticosteroids which causes them to compete aggressively and exclude black-heads from nesting cavities. As the proportion of red-heads increases, so does the cost of competition with other, red-headed birds. This reduces the resources available to red-heads for parenting, so that black-heads, who invest more resources in parenting, raise more successful broods. In wild populations the two types persist alongside one another at around three black to one red and with occasional gene flow between them.

Fig 2. Picture of Gouldian Finches

What are the functions of these behavioral traits in Gouldian finches? We can apply Defn. 7 to determine the Proper function of the Hawk-ish phenotype of red-headed finches:

Consider a population of finches, some of whom have elevated production of testosterone and corticosteroids (Elevated). These organisms are descendants of finches with Elevated and their character states are homologous. In this population, Elevated has defeating rivals for nesting cavities as a direct Proper function because:

H1. In some ancestral populations there was variation in production of testosterone and corticosteroids (Elevated vs. Reduced)
H2. Elevated caused some ancestral Elevated individuals to defeat rivals for nesting cavities with higher probability than individuals with Reduced
H3. Defeating rivals for nesting cavities was causally responsible for some ancestral Elevated individuals having greater reproductive output than they would have had if Elevated had been changed to Reduced
H4. The frequency of Elevated in the current population is explained by selection for Elevated in these ancestral populations in virtue of the increase in fitness accruing to Elevated individuals from defeating rivals for nesting cavities more frequently than Reduced individuals.

We can also apply Defn. 7. To determine the proper function of the Dove-ish phenotype of black-headed finches:

Consider a population of finches, some of whom have reduced production of testosterone and corticosteroids (Reduced). These organisms are descendants of finches with Reduced and their character states are homologous. In this population, Reduced has conserving resources for parental investment as a direct Proper function because:

D1. In some ancestral populations there was variation in production of testosterone and corticosteroids (Elevated vs. Reduced)
D2. Reduced caused some ancestral Reduced individuals to conserve resources for parental investment with higher probability than individuals with Elevated
D3. Conserving resources for parental investment was causally responsible for some ancestral Reduced individuals having greater reproductive output than they would have had if Reduced had been changed to Elevated
D4. The frequency of Reduced in the current population is explained by selection for Reduced in these ancestral populations in virtue of the increase in fitness accruing to Reduced individuals from conserving resources for parental investment more frequently than Elevated individuals.

Definition 7 assigns a proper function in both cases, and these function ascriptions also seem to be in the spirit of the six earlier definitions. But the claim that the Proper function of each trait explains why the trait exists is highly problematic.

Imagine the scenario where a population of red-headed individuals is invaded by a small number of black-headed individuals. The proportion of red-heads will decline until it reaches equilibrium. Throughout this process red-heads have lower reproductive fitness than black-heads. So Defn 7 will not apply, because clause 7.3 is not fulfilled. It was never the case in the actual evolution of the observed ratio of red-heads that Elevated (red-head) had higher reproductive output than Reduced (black-head). So it is not the Proper Function of Elevated to defeat rivals for nesting cavities. But nothing of significance in the biology corresponds to this dramatic loss of Proper function. The model of the evolutionary forces acting in the population is unchanged and gives the same explanation of why around 25% Elevated individuals exist in the
population at equilibrium. But if the population reaches equilibrium by this trajectory, the existence of Elevated is not explained by Elevated performing its Proper function.

We suggest that even when the definition of Proper function succeeds in assigning a function to Elevated, the existence and proportion of this trait in a population is not explained by its Proper function. It was the desire to fit such cases into the theory of Proper functions that led Griffiths (1993) to offer Defn. 3 (above). Griffiths fitted these cases by weakening the sense in which Proper functions ‘explain’ traits: the Proper function is merely an explanatorily relevant factor; an explanation that omits it is not a complete explanation. In the finch case, the positive effects on fitness of Elevated is certainly an explanatorily relevant factor - if you change it then the outcome changes. But the key to understanding why we observe red- (and ipso facto black-) headed finches is the variable effect of traits on fitness.

The fact that red-heads consistently defeat black-heads in competition for nesting cavities is part of an evolutionary explanation of why red-heads exist. But when removed from that broader explanation, this fact alone does not explain the existence and proportion of red-heads. It is simply not true that defeating rivals for nesting cavities enhances fitness - it only does so when the population is <25% red-headed. The true explanation of the proportion of red- and black-heads is how the relative fitness of each type changes as the composition of the population changes. It is not even true that Proper function explains why red-heads at least exist (in whatever proportion). That too will depend on the rest of the model. In this particular case, models reveal a “population fitness valley, [where] hawks are competent enough to take over nest sites from doves and to breed relatively successfully, but are still not successful enough to maintain the productivity of the population in the absence of the healthy population growth that doves provide.” (Kokko et al. 2014, p. 4). In these scenarios, the result of the ‘success’ of red-heads in defeating their rivals for nesting cavities is that both types go extinct, first black-heads and then red-heads.

The selected effects theory is built upon the following simple, intuitive picture of evolution. An organism with character $C_i$ has Proper function $F$ because performing $F$ gave it a reproductive advantage over an organism with character $C_j$. Repeated interaction between $C_i$ and $C_j$ individuals, each interaction resulting in a reproductive advantage for $C_i$, explains the existence of $C_i$. We have shown that this simple, intuitive picture often does not hold: interactions between $C_i$ and $C_j$ individuals change the environment, and this change in environment feeds back and alters the payoff for $C_i$, such that $C_i$’s reproductive advantage may vanish or even reverse. But before we sound the death knell of Proper function, we must consider an extension of selected effects theory that has been formulated specifically to deal with certain issues that arise in heterogeneous environments: Millikan’s concept of ‘Normal’ environment. In the following section, we introduce ‘Normal’ environments and show why they are insufficient to solve the problem we have highlighted.
3.3 Normal Environments

The idea of the Normal (big ‘N’) is at the heart of Millikan’s theory of Proper function: “I capitalize Normal — to distinguish it from normal in the sense of average.” (Millikan 1984, p. 34, italics in original). To be Normal is to be the way things were on actual, historical occasions on which a trait successfully performed its functions. Normal conditions are “conditions in which [the trait] has historically been when it actually performed F - these conditions being uniform over as large a number of historical cases as possible.” (Millikan 1984, p. 33). A Normal explanation is one that explains how F is performed under these Normal conditions. A Normal character is a character state that features in a Normal explanation (Appendix 1).

It is important to understand how Millikan’s Normal differs from conventional ideas of normal (small ‘n’) environment or historical environment in biology. The normal environment in the conventional sense includes both environments in which the trait was historically successful and environments in which it was historically unsuccessful - it is both normal for rabbits to escape predators and normal for rabbits to be eaten. But Millikan’s Normal environment is the set of environmental conditions which are needed for the trait to perform its function successfully in just the way it performed it in the evolutionary past. Only environments in which the rabbit escapes, and moreover escapes in the way it did in the past, are Normal environments in which the rabbit’s anti-predator adaptations perform their Proper functions in accordance with a Normal explanation.

Millikan developed the concept of Normal for two reasons: (i) to distinguish a trait performing its function as it did in the evolutionary past (Proper functioning) from the trait performing the same function merely adventitiously; and (ii) to make sense of traits malfunctioning in their normal (small ‘n’) environments. To understand what is meant by (i) consider a male Superb Fairy Wren who has an extra-pair copulation because the female wren remembers that he was the first local male to turn blue in spring. Here the blue trait performs its function in accordance with a Normal explanation. But suppose a male wren has an extra-pair copulation because the owner of an aviary picks the bluest wrens in an effort to breed males who are constitutively blue. According to Millikan, the blue trait does not have its effect as a Proper function in this case because it does not produce the effect in accordance with a Normal explanation.7 To explain what is meant by (ii), Millikan offers the example of a sperm unable to perform its Proper function of fertilizing an egg:

“Consider again how few sperm have historically managed to realize any but the most immediate functions (say, swimming) proper to them. This is because very few sperm

---

7 Millikan emphasises that a trait can function Normally in a narrow environment containing all the ‘proximate’ Normal conditions for successful performance of its function despite the fact that more distal conditions are abNormal. For example, a spermatozoon may propel itself and bind to an egg Normally despite the fact that the prostate secretions, vaginal secretions and egg have been provided artificially. In that case, Millikan says that the sperm functions in accordance with a ‘proximate Normal explanation’. But the full range of ‘distal’ Normal conditions are needed to connect this ‘proximate’ function to reproductive output - to give a Normal explanation of reproductive output.
So Normal conditions are the conditions which allowed the trait to successfully perform its function in a non-accidental way in actual, historical instances where the function was performed.

Millikan says that Normal conditions “are the conditions to which the device that performs the proper function is biologically adapted.” (Millikan 1984, p. 34). But this is simply incorrect. Whether a trait succeeds in evolutionary competition will depend on its relative fitness in the expected environment, a weighted average of environments in which the trait occurs, not its relative fitness in the ‘big-N’ Normal environment(s). Traits are adapted to their expected environments,\(^8\) which include both Normal and abNormal environments in Millikan’s senses.

The very idea of Normal explanation implicitly divides the environment into Normal patches, where the character state is advantageous, and abNormal patches where the character state is not advantageous. The abNormal patches are ignored in the Normal explanation of how the character state affects fitness and hence in the Proper function explanation of why we observe that state in the target population. But actual evolutionary models of selection in patchy environments must consider how the trait performs in both advantageous and disadvantageous patches, as we will now explain using the example of bet-hedging in heterogeneous environments.

### 3.4 Heterogeneous environments and bet hedging

Evolution in patchy environments, whether spatial or temporal, does not involve an evolutionary response to a single kind of patch irrespective of what is happening in other patches (Brandon 1990). Organisms evolve to deal with the heterogeneity inherent in the environment, responding to a probability distribution over possible environments with which they may have to cope. One well-known response to patchy environments is a ‘bet-hedging’ strategy.

A trait embodies a bet hedging strategy if it minimises variance in fitness such that the trait’s arithmetic (additive) fitness decreases. If environmental conditions are sufficiently variable, causing temporal fluctuation in a trait's fitness, then a mutant trait with a lower arithmetic mean fitness may spread on account of its geometric (multiplicative) mean fitness being larger than the resident’s (Gillespie 1974). To illustrate, consider a simple example of conservative bet-hedging (‘play it safe’). Imagine two environments - one good and the other bad - that occur with equal probability. Say an organism that does not employ bet-hedging has a relative reproductive output of 1 in the good environment and a relative reproductive output of 0.5 in the bad environment. That organism’s arithmetic mean reproductive output is \(1 * 0.5 + 0.5 * 0.5 = 0.75\)

---

8 In the simplest case - sometimes a more complex statistic than the expected value is needed (Sober 2001)
and its geometric mean reproductive output is $\sqrt{1 \times 0.5} = 0.707$. Consider now an organism that employs conservative bet-hedging that has a reproductive output of 0.71 in both good and bad environments. Its arithmetic and geometric fitnesses are both 0.71. Since natural selection is a multiplicative process, the bet-hedging organism has a higher expected number of offspring, despite having a lower arithmetic expectation.

A canonical biological example of bet hedging is seed dormancy in annual plants. For many annual plants, the environmental conditions for seedling establishment and reproduction can vary widely from year to year. Although a strategy in which all seeds germinate the following year is optimal when conditions are predictably good, such a strategy risks catastrophic failure if the following year is a drought. To get around this, some annual plants evolve a diversified bet-hedging strategy (‘don’t put all your eggs in one basket’) in which they produce seeds that germinate in different years (a so-called "seed bank" (Childs et al. 2010)). Let us consider the form of a Proper function explanation of bet hedging in plants.

We assume that each year the environment is either Good or Drought. All seeds that germinate in a Good year survive, but all seeds that germinate in a Drought year die. As per Defn 7, let the character C be Germ_Pattern, of which we assume there are only two variants: a trait C_i (denoted Variable_Germ) that causes half the seeds to germinate in the following year and half the year after, and a trait C_j (denoted Consistent_Germ) that causes all seeds to germinate the following year. Furthermore, we assume that the effect F of Variable_Germ is to decrease the proportion of seeds that germinate in Drought (denoted Reduce_Drought_Germ), which it does more efficiently than Consistent_Germ. We can now fill out Defn. 7.

Consider a population of organisms some of whom have Variable_Germ. These organisms are descendants of organisms with Variable_Germ and their character states are homologous. In this population, Variable_Germ has effect Reduce_Drought_Germ as a direct Proper function if and only if:

1. In some ancestral populations there was variation in Germ_Pattern
2. Having Variable_Germ caused some ancestral Variable_Germ individuals to Reduce_Drought_Germ with higher probability than Consistent_Germ individuals.
3. Reduce_Drought_Germ was causally responsible for some ancestral Variable_Germ individuals having greater reproductive output than they would have had if Variable_Germ had been changed to Consistent_Germ in that ancestral population.
4. The frequency of Variable_Germ in the current population is explained by selection for Variable_Germ in these ancestral populations in virtue of the increase in fitness accruing to Variable_Germ individuals from Reducing_Drought_Germ more frequently than Consistent_Germ individuals.

This seems to be a legitimate explanation. After all, if somebody were to ask “why do this plant’s seeds not germinate every year?”, a reasonable explanation is “because it is hedging its bets against the possibility of catastrophic drought conditions”. It does not run into the problem we
saw with coevolution and frequency-dependence because there is no trait-environment feedback modulating the trait’s effect (i.e. the fitness of $C_i$ is independent of $C_i$’s frequency).

The problems arise when we try to apply a Normal explanation. The diversified bet-hedging of the tree works by limiting losses in Drought years so that ultimately more seeds can germinate in Good years. If all seeds die once a Drought hits, then there are no seeds left to germinate in the next Good year. Crucially, to understand whether the bet-hedging trait will spread, we must account for both Drought and Good environments and their respective likelihoods (the ‘small n’ normal environment). Indeed, our assignment of Proper function earlier implicitly accounts for both environments (e.g. the truth of clause 3 relies on our assumption that both Good and Drought conditions exist). Milikan’s concept of Normality, however, isolates a single environment and attempts to explain a trait’s existence solely based on its performance in this Normal environment. As a result, if we tried to assign a Normal environment to a bet-hedging trait, we would only be capturing part of the picture.

In this example, we believe that Millikan would instead apply the Normal environment - and thus the Normal explanation - to individual seeds. For example, we might say that the Normal environment for a seed is Good, and that the Normal explanation is that the Proper function of the seed is to germinate in the Good environment. But this omits the key piece of explanatory information behind the bet-hedging strategy - the reason that a seed produced by a tree with the bet-hedging trait can germinate in a Good environment is because some seeds did not germinate during Drought years. We could try reversing the explanation and assigning Drought as the Normal environment. In this case, the Proper function of the seed would be to not germinate in the Drought environment, but this does not help because it omits explanatory information about germinating in Good years. Ultimately, to make sense of traits that evolve in heterogeneous environments, we must invoke both the Normal and abNormal environments (i.e., we must consider the ‘small n’ normal environment).

3.5 When and Why Proper functions do not Explain

The distinctive appeal of the selected effects theory of Proper functions is that Proper functions so defined are supposed to explain the traits that bear them. Male fairy wrens have blue heads because those heads attract mates. Red-head Gouldian finches have elevated testosterone and corticosteroids because this allows them to win fights over nest sites. In a Proper function explanation, the existence and/or prevalence of the trait ($C_i$) is explained by pointing out one or more kinds of interaction between trait and environment in which the fitness of $C_i$ is higher than it would have been had $C_i$ been replaced by some alternative form $C_j$. It is the iterated occurrence of these interactions that explains why the trait has been as successful as it has been. It is evident that these must be the explanatory facts that constitute a Proper function explanation, because these are the only facts provided by knowing the Proper function (Defn 7, clauses 1-3). Clause 4 cannot contain any extra explanatory facts, since clause 4 essentially says ‘and the facts cited in 1-3 explain why the observed trait evolved by natural selection’. Of course, as noted above (fn. 4) it is open to the Proper function theorist to restrict the theory to cases where clauses 1-3 contain all the important explanatory information. We would regard this
as an admission of defeat, since traits with non-trivial evolutionary dynamics would not have Proper functions.

So why do Proper functions not explain in the cases we have highlighted? First, because whether $C_i$ evolves (or what frequency it reaches) often depends on considerations of relative fitnesses in environments where the fitness of $C_i$ is less than the fitness of an alternative trait $C_j$. These are ignored in the definition of Proper function, which focuses exclusively on situations where the fitness of $C_i$ is consistently greater than the fitness of $C_j$. Second, because when environmental heterogeneity is produced endogenously, as in frequency-dependent selection, whether $C_i$ evolves (or what frequency it reaches) depends not only on the relative fitnesses of $C_i$ and its alternatives across a range of environments, but also on many factors such as the starting ratios, the order in which alternative character states arise, and more generally, on the evolutionary dynamics.

We must emphasise that we are not merely pointing out that Proper function explanations are incomplete. In the cases we have described, the information contained in a Proper function ascription omits information needed to predict the outcome of the selection process. But, more significantly, that additional information cannot reasonably be regarded as background conditions or a ceteris paribus clause. In these cases, we have a thorough evolutionary explanation, but the Proper function theory extracts just the information that would be of explanatory relevance if the key phenomenon in that explanation - such as environmental heterogeneity or frequency-dependence - were not present. The resulting ‘explanation’ is not merely incomplete, it is wrong. In these cases, the mere fact that a character state has repeatedly faced environments in which it was fitter than some alternative state is, if true, merely one fragment of explanatorily relevant information. It is not ‘why the trait evolved’ as Proper function theorists have always insisted.

We do not deny that Proper functions can explain traits when none of these evolutionary phenomena are present. But in the next section we explain why this is cold comfort for Proper function theorists.

4. Implications

The selected effects account of Proper function seems straightforward when we focus on the simple, intuitive case of a trait that solves a fixed problem better than its competitors. Problems arise when phenomena such as environmental heterogeneity and trait-environment interactions that feedback on the fitness of the trait are present. We acknowledge that such feedback may not always cause problems for Proper function explanations. For example, negative frequency-dependent interactions might be weak, such that the fitness of trait $C_i$ diminishes as $C_i$ spreads but nevertheless always remains higher than the fitness of alternative $C_j$ until $C_i$ has displaced $C_j$. Likewise, positive frequency-dependence dynamics can mean that trait-environment feedback causes $C_i$’s fitness to actually increase relative to $C_j$ as $C_i$ spreads. For the purposes of Proper function, these cases (and possibly others) are similar enough to the intuitive case of
a trait that solves a fixed problem better than its competitors as to not cause problems. But, as we have shown, there are numerous examples of evolutionary dynamics that do not correspond even approximately to the intuitive picture.

Proper function theorists have long recognized that the domain in which the theory is applicable depends on the nature of the selection process: “Differences between biological and artifact functions arise because of the different nature of the selection process involved in the two cases” (Neander 1991a). We have taken this a step further by showing that the theory does not even generalize to all cases of natural selection. These problem cases cannot be dismissed as unusual or peripheral examples of natural selection. Frequency dependent selection, environmental heterogeneity and more generally the coevolution of traits and their environments (which may include other organisms of the same or different species) are ubiquitous (Yoshimura and Clark 1993; Asmussen et al. 2004; Brockhurst et al. 2014). For traits evolving under these dynamics, we can often define Proper functions, but these Proper functions fail to explain the trait. The problem is that the theory is designed to extract from any evolutionary scenario just that information that would be explanatorily relevant if the focal trait was evolving without either exogenous or endogenous environmental heterogeneity. So when these phenomena are present, the theory of Proper function omits by design the key explanatory information in the evolutionary scenario. The explanation it produces is wrong - by design.

4.1 Signalling as a problem for Proper function

Importantly, feedback between trait and environment applies with particular force for the evolution of traits that are of the greatest interest to Proper function theorists. Millikan’s primary aim in developing her theory of Proper function was to ascribe functions to representations - phenotypes used in signalling and communication. Explaining representational states is also the central aim of proper function theorists Neander and Shea (Shea 2007; Neander 2017). But the evolution of representational states is the coevolutionary process par excellence.

According to influential accounts in biology (Maynard-Smith and Harper 2003) and philosophy (Scott-Phillips 2008; Skyrms 2010), the best models of how communication evolves are inescapably co-evolutionary. Information-bearing signals might evolve between different species (such as in the stotting displays of Thompson’s Gazelles, to their predators), within a species (such as with sexual signalling (Zahavi 1975; Grafen 1990)) and even within the same organism, as in gene regulation (Calcott et al. 2018). Since Lewis (1969), such models have also been used to explain the emergence of language and other conventions (Huttegger 2007). Similar models are used regardless of the specific systems they are applied to, generally utilizing evolutionary game theory and a sender-receiver framework (Skyrms 2010; Huttegger and Zolman 2013). In these models there is a sender agent which is sensitive to some state of the world, and a receiver agent which is able to observe the sender. Paradigmatic, informational communication occurs when the sender conditionalizes an observable behaviour on states of

---

9 It will also be important to assess the potential of the theory of Proper function in light of niche construction theory (Odling-Smee et al. 2003) but here we appeal only to issues that face Proper function theorists within the most conventional, neo-Darwinist picture of evolution.
the world (perhaps including internal states), and the receiver conditionalizes their behaviour on
the sender’s observed behaviour/signal. These conditional behaviours (would-be signalling
strategies) can be stable as a pairing (i.e. be a signalling equilibrium) if the agents both receive
a better payoff from the effects of their combined behaviours than they would from other
behaviours they have available. And these behaviors can evolve if they are repeated under
familiar evolutionary conditions where payoffs can be interpreted as fitnesses: i.e. if senders and
receivers mostly retain their strategies but update them to follow a fitness gradient which rises
from their initial strategies to peak at the signalling equilibrium\(^\text{10}\). In such cases, sender &
receiver signalling strategies (whether they be stotting, peacock tails) should be amenable to
explanation in terms of Proper function.

Signalling models are co-evolutionary because the fitness of a sender strategy at a specific time
depends on the prevalence of complementary receiver strategies, and vice versa. There is no
advantage in sending a signal if no-one is paying any attention to it, or if the responses are
adverse. Likewise, there is no advantage in being poised for signals that are not forthcoming or
are poorly correlated with reality. This makes the evolution of signalling frequency dependent.
As in some of the examples discussed above, the effective selective environment includes the
opposing communication partner’s evolving traits and is therefore not static, and this is
endogenous to the co-evolutionary relationship between sender and receiver strategies. Indeed,
in evolutionary simulations, the likelihood of settling into a signalling equilibrium is highly
sensitive to the initial mix of strategies, especially as modelling idealisations are relaxed
(Pawlowski 2008; Hofbauer and Huttegger 2008; Huttegger and Zollman 2013), though the
effects of this are unpredictable and sometimes surprising (Brusse and Bruner 2017). In all
cases, the sender-receiver pairing that evolves is explained by the dynamic co-evolution of the
constituent strategy traits, and their effective fitness can fluctuate wildly in response to each
other.

The Proper function theorist might be tempted to assume that frequency dependent effects will
‘even out’ as populations begin to converge on a compatible sender-receiver strategy pairing.
However there is one variety of signalling equilibria, so-called hybrid equilibria, where this is not
at all the case. In a hybrid equilibrium, populations of sender and receiver strategies do not
fixate or settle into fixed proportions of signalling and non-signalling strategies, but instead orbit
around a weakly stable equilibrium point, in a manner similar to hawk-dove games. These
‘partially honest’ dynamic equilibria are traditionally under-examined but are not just oddities:
they have been demonstrated in simple handicap signalling games over wide, varied, and
realistic parameter spaces (Zollman et al. 2013; Bruner et al. 2017). As (Kane and Zollman
2015) attest, they may have considerable evolutionary significance and application both as
bridging regimes to more classical communicative equilibria, and as communicative equilibria in
their own right.

---

\(^{10}\) This describes models based on the so-called replicator dynamics. Models which instead appeal to
learning dynamics or other update rules have broadly analogous commitments and outcomes, such that
the replicator dynamics is often used to model real-world target systems even where biological
reproduction is clearly not the mechanism by which strategies are updated.
It is not possible to explain the communicative behaviours of agents in hybrid equilibria via Proper function. The rest points that hybrid equilibria orbit around depend on the cost structure and other parameters of the signalling game. But the orbiting dynamics of hybrid equilibria mean that there is an even split over time between communicative strategies having positive and negative relative fitness - they decline, then increase again. As before, we run into trouble if we attempt to gerrymander a Normal environment out of a variable selective environment like this. If we aggregate many orbital cycles into a Normal environment (e.g. by extending out ‘evolutionary recent’ history to incorporate them), then we get an inferred averaged environment which would explain an idealised strategy mix. But this is not the actual strategy at any given time, and treating the environment as a given ignores the co-evolutionary relationship that explains its variation, which seems ad hoc. Alternatively, the relevant time-frame might be taken to be much shorter, referencing the fitness of strategies against a more transient Normal environment. But this would see signalling strategies (and hence the Normal) oscillating wildly. Neither description seems right. In any case, we cannot say (with reference to our definition 7) that there is an increase in fitness accruing to individuals from performing the signalling function more frequently than individuals with alternative character states - the regime is alternating.

The issues confronting Proper function accounts of signalling and communication reveal a key problem with the existing literature on Proper function. Proper functions are introduced using simple, intuitive examples of selection in a constant environment, like the zebra stripes in section 3.1. In such cases it is plausible that there are Proper function(s) that explain the existence and prevalence of phenotypes. But the main use of the theory is in application to phenotypes whose evolution violates all the assumptions that make that claim plausible. Whether these phenotypes have Proper function(s) that explain their existence and prevalence is an open question.

4.2 Where now for Proper function theories?

Proper function theorists will be tempted to suppose that even if it is not universally applicable, the current, simple theory nevertheless remains a useful abstraction for thinking about biological functions. We disagree for two reasons. First, the central appeal of the selected effects account is its claim that only it can explain why biological traits exist. It will not do for it to explain a handful of traits but provide only a fragment of explanatory information for all the others. Other, rival theories such as the ‘organisation’ theory of function also meet that very weak criterion (Mossio et al. 2009; Garson 2016). Proper function theorists have strongly criticised organisational theories because the ‘organisational function’ of a trait does not explain the existence and prevalence of that trait. But the ‘organisational function’ of a trait is how the trait contributes to the ability of an organism to maintain and replicate itself, which is surely of some explanatory relevance to why the trait exists.

Second, and more importantly, an abstraction is only useful when it retains the elements that influence important dynamics. In mathematical modelling a well-designed model abstracts away extraneous aspects while retaining those that meaningfully influence the dynamics of the
process under study. If a modeller removes important components then their inferences will go astray. Consider a real-life biological example. In the early 90s, models were used to support the hypothesis that uniparental (maternal) inheritance of mitochondria drove the evolution of mating types (primitive sex cells) (Hurst and Hamilton 1992; Hutson and Law 1993). Twenty-years later, another paper came to the opposite conclusion (Hadjivasiliou et al. 2013). What happened? The early models made a simplifying assumption that inadvertently removed frequency-dependent dynamics from the models, altering the evolutionary dynamics of the traits, and leading to an incorrect inference (Hadjivasiliou et al. 2013). Considerations such as these are precisely the reason that evolutionary biology is founded on quantitative principles. Models serve as a check on intuition, as a safeguard against minor conceptual errors that can easily lead our inferences astray. So the shortcomings of current definitions of Proper function are just the kind that make for unreliable, potentially misleading abstractions.

We do not want to simply reject evolutionary or ‘teleonomic’ approaches to function. Confronting the complexities of real population biology might allow for the development of a more sophisticated theory. But the development of such a theory will need to overcome some deep and demanding problems.

The problems we have identified for theories of Proper functions reflect serious tensions between our current understanding of evolutionary dynamics and the intuitive idea that natural selection produces adaptations whose purpose is to increase fitness in some specific way. In a model of a large population in which there is one locus with two alleles, the process of natural selection, assuming it acts solely on viability and is frequency independent, can only increase the mean fitness of the population, or at worst leave it unchanged (Wright 1937). Increase in mean fitness is considered a proxy for adaptation in this literature. This apparently vindicates natural selection as a process that tends to maximize fitness, or as an improvement process (Rice 2004). However, the assumptions of this model are very restrictive. To reiterate, there are only two alleles at one locus, the selection occurs on the viability of individuals, and selection is frequency independent. Although this result has been generalised for one locus with any number of alleles (Kingman 1961), relax any other assumption and the conclusion that natural selection necessarily leads to increased mean fitness no longer holds. Different scenarios under which these assumptions do not hold have been extensively studied by population geneticists (see Ewens 2004) and many have reached the conclusion that the link between natural selection and adaptation may be much weaker than is usually supposed. Negative frequency dependent selection, an example discussed at length above (Gouldian finches), is a case in point. If fitness is not stable across generations because of frequency dependence or other processes, then mean fitness in a population, a common proxy for adaptation, will not necessarily increase (Rice 2004, p. 14).

The weakness of the link between natural selection and adaptation has left some evolutionary biologists unsatisfied. This is particularly true in behavioral ecology, where it is classically assumed that natural selection can be glossed as the maximisation of inclusive fitness and organisms studied using adaptationist reasoning that reflects this assumption. If the link between selection and adaptation is strong under idealised assumptions but breaks down when
these are violated, behavioral ecologists are faced with a serious challenge. Alan Grafen developed his ‘formal Darwinism’ project with the explicit aim of meeting this challenge (Grafen 2014).\textsuperscript{11} Despite the ingenuity of this work, many, including the philosophers who have examined this controversy in detail (Birch 2016; Okasha 2018), believe that he does not succeed or only partly succeeds. So the problems of relating natural selection, optimization and adaptation remain unresolved. It is these problems that a biologically plausible theory of Proper function will need to confront. This is not an easy challenge and will, at the very least, require an integration of philosophical work on function with the philosophy of population biology.

\textsuperscript{11} For a philosophical introduction of the project, see the 2014 special issue in Biology and Philosophy.
### Appendix: Millikan’s technical terminology and conventional biological terminology

<table>
<thead>
<tr>
<th>Millikan's term</th>
<th>Our translation</th>
<th>Illustrative examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>First-order reproductively</td>
<td>Homologue of genes or genotypes. In molecular biology, orthologs, paralogs and</td>
<td>16S rRNA gene, clonal line of asexual organisms</td>
</tr>
<tr>
<td>established family</td>
<td>xenologs will all count as 'families' of genes.</td>
<td></td>
</tr>
<tr>
<td>Higher-order reproductively</td>
<td>Homologue of phenotypes, including behavioural phenotypes.</td>
<td>16S rRNA, femur, courtship display of the Great Crested Grebe</td>
</tr>
<tr>
<td>established family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Device</td>
<td>A mechanism with a heterocatalytic function (it makes something that is not a</td>
<td>Machinery of DNA transcription and translation, stem cell, brains ('making behaviours')</td>
</tr>
<tr>
<td></td>
<td>copy of itself)</td>
<td></td>
</tr>
<tr>
<td>Proper function</td>
<td>An effect for which a character is an adaptation (where 'adaptation' includes</td>
<td>Proper function of 16SrRNA gene is to code for 16SrRNA; proper functions of femur</td>
</tr>
<tr>
<td></td>
<td>exaptations which have been subject to purifying selection in the recent past)</td>
<td>include strengthening hindlimb, haematopoiesis, etc; proper function of courtship</td>
</tr>
<tr>
<td></td>
<td></td>
<td>display is to create pair bond</td>
</tr>
<tr>
<td>Normal</td>
<td>How something was historically when the function was successfully performed, as</td>
<td>Normal environment for a spermatazoon to perform the function of swimming in accordance</td>
</tr>
<tr>
<td></td>
<td>in 'Normal environment', 'Normal character', 'Normal explanation'. See text for</td>
<td>with a Normal explanation (see Millikan 1984, p. 34) includes acidic vaginal fluid,</td>
</tr>
<tr>
<td></td>
<td>discussion.</td>
<td>fibryno lysin enzymes, etc</td>
</tr>
</tbody>
</table>
References


Neander K (1983) Abnormal Psychobiology. LaTrobe University


