Neo-Paleyan Biology

Tim Lewens

University of Cambridge

Department of History and Philosophy of Science

Free School Lane

Cambridge CB2 3RH

Email: tml1000@cam.ac.uk

Abstract

There is a 'Neo-Paleyan' tradition in British evolutionary theorising, which began with Darwin and continues to the present day. This tradition conceives of adaptation in terms of design, and it often puts natural selection in the role of an ersatz designer. There are significant disanalogies between Paleyan conceptions of design and modern conceptions of adaptation and selection, which help to explain why the neo-Paleyan programme is sometimes treated with hostility. These general disanalogies do not suffice to dismiss the most interesting forms of recent neo-Paleyanism, which draw on theoretical principles such as Fisher's Fundamental Theorem to ground a general approach to what we can call (following Grafen) the 'criterion' of evolutionary design. It is important to distinguish between justifications of this 'criterion' and justifications of approaches to nature which presuppose that natural selection produces good designs.

Keywords: Adaptation, Andy Gardner, Alan Grafen, Design, Fisher's Fundamental Theorem, William Paley

Funding

This work was funded by a grant from the John Templeton Foundation. I am also grateful to the Fondation Maison des Sciences de l'Homme, Paris, for support during the completion of this project.

Acknowledgements

Earlier versions of this paper were presented in Copenhagen, Oulu and Paris. I am grateful to audiences there, and to the referees from this journal for comments. I am especially grateful to Jean-Baptiste Grodwohl for remarkably detailed and penetrating remarks on the submitted manuscript, and to Andy Gardner, Jonathan Birch and Samir Okasha for additional suggestions for improvements and clarifications.

1. Neo-Paleyan Biology

Richard Dawkins once described himself as a 'neo-Paleyist' (1998: 16). He is not the only prominent evolutionary biologist to have praised the natural theologian William Paley (1743-1805), in spite of the very same biologists opposing appeals to intelligent design in the explanation of the form of plant and animal species. The typical neo-Paleyan package gives Paley credit for having identified the central problem for the natural historian, while adding that he entirely mistook the true process that answers that problem.

This article establishes the existence of this neo-Paleyan tradition, clarifies some of its most common themes, and warns against potential misunderstandings of it. It moves on to scrutinise just one recent incarnation of neo-Paleyanism in detail, namely the design-based approach of Gardner and Grafen. I examine a series of arguments suggested by Gardner (2009, 2017), which suggest various links between the notion of adaptation and the notion of design. I then address the powerful, and ongoing, efforts of Grafen and Gardner to draw links between design-thinking and Fisher's Fundamental Theorem. (In the interests of expository simplicity, I will not discuss here Grafen's related, and more technically demanding, project of justifying the organism-as-maximising-agent approach via the Price Equation (see Grafen 2014 inter alia), except insofar as it overlaps with Grafen and Gardner's invocations of the Fundamental Theorem.)

I argue that we need to distinguish two roles for the Fundamental Theorem, or indeed, for any attempt to formalise a design-based approach to evolution. We might try to use the Theorem to justify an expectation of finding good design in nature; or we might try to use it to spell out what Grafen (2015: 1) has called the 'criterion' of evolutionary design, i.e. the quantity most generally favoured by natural selection processes. I argue (in a manner that is intended to cohere with Grafen and Gardner's own comments) that while the Fundamental Theorem may be able to achieve the second task, it does not give us a demonstration that natural selection has a general tendency to raise design quality. It appears, though, that some neo-Paleyans within behavioural ecology and evolutionary psychology presuppose this stronger conception of natural selection's link to the production of design.

2. *Natural Theology*

The evolutionary biologists whom I describe as 'neo-Paleyans' typically assert that the primary explanatory task of evolutionary theory is to account for the existence of adaptations. They characterise adaptations, in explicit reference to the earlier natural theological tradition, as traits that give the appearance of good design. As Gardner and Welch put it: 'The cardinal problem of biology is to explain the process and purpose of adaptation, i.e. the apparent design of the living world' (2011: 1).

The upshot of this combination of views, at least in Richard Dawkins's case, is that he ends up giving an agent-dependent definition of adaptation. He links the notion of adaptation to the notion of good design, and the notion of good design to that which an intelligent and proficient engineer—i.e. an agent—might want to construct:

> We may say that a living body or organ is well designed if it has attributes that an intelligent and knowledgeable engineer might have built in order to achieve some

> sensible purpose, such as flying, swimming, seeing, eating, reproducing, or more generally promoting the survival and reproduction of the organism's genes. (1986: 21)

Neo-Paleyans argue that it is natural selection that explains this good design. Since some also argue that it is *only* natural selection that can explain good design, they sometimes defend an updated analogue of Paley's argument from design (see Lauder 1996 for discussion). Good design, they say, is diagnostic of the action of natural selection. Pinker and Bloom, for example, tell us that:

> Evolutionary theory offers clear criteria for when a trait should be attributed to natural selection: complex design for some function, and the absence of alternative processes capable of explaining such complexity. (Pinker and Bloom 1990: 707)

This section aims merely to give further evidence that this neo-Paleyan tradition exists.

Paley

Paley's *Natural Theology* (1802) offers a large catalogue of what he tends to call 'mechanisms' (Gillespie 1990). A mechanism, for Paley, is any assemblage of orchestrated parts in which we see, just as in a watch, 'the suitableness of these parts to one another; first, in the succession and order in which they act; and, secondly, with a view to the effect finally produced' (1802: 140). Such mechanism, says Paley, cannot be explained by chance, only by what he calls 'contrivance': 'What does chance ever do for us?', he asks: 'Amongst inanimate substances, a clod, a pebble, a liquid drop might be; but never was a watch, a telescope, an organized body of any kind, answering a valuable purpose by a complicated mechanism, the effect of chance' (ibid.: 38).

Paley's comments on the woodpecker are typical:

> The woodpecker lives chiefly upon insects, lodged in the bodies of decayed or decaying trees…The bird, having exposed the retreats of the insects by the assistance of its bill, with a motion inconceivably quick launches out at them this long tongue; transfixes them upon the barbed needle at the end of it; and this draws its prey within its mouth. If this be not mechanism, what is? Should it be said, that, by continual endeavours to shoot out the tongue to the stretch, the woodpecker species may by degrees have lengthened the organ itself, beyond that of other birds, what account can be given of its form; of its tip? How, in particular, did it get its barbs, its dentition? These barbs, in my opinion, wherever they occur, are decisive proofs of mechanical contrivance. (Paley 1802: 133-4)

It is this combination of marvel at the intricate design of organic nature, coupled to a sense that design must be explained by some process other than chance, that infuses generations of neo-Paleyans who follow.

Darwin

Charles Darwin is the first of our neo-Paleyans. He eventually satisfied himself that 'The old argument from design in nature, as given by Paley, which formerly seemed to me so conclusive, fails, now that the law of natural selection has been discovered' (Darwin 2002: 50). Even so, Paley's influence is palpable in Darwin's work. Darwin remarked in his *Autobiography* that during his time as an undergraduate in Cambridge, 'The logic of [Paley's *Evidences of Christianity*] and as I may add of his *Natural Theology* gave me as much delight as did Euclid' (ibid.). At the very beginning of the *Origin* he makes it clear that any transformationist theory—any theory, that is, which claims that the species we find around the planet are modified descendants of some small number of common ancestors—would be incomplete unless it were able to answer what we might call Paley's Question. This is the question of how, 'the innumerable species inhabiting this world have been modified, so as to acquire that perfection of structure and coadaptation which most justly excites our admiration' (Darwin 1859: 3).

At this point Darwin, like Paley, turns to the woodpecker. By itself, the hypothesis of common ancestry contains nothing that might explain 'the structure…of the woodpecker, with its feet, tail, beak, and tongue so admirably adapted to catch insects under the bark of trees' (1859: 3). Like Paley, the broad question that Darwin focuses his explanatory efforts on is to understand, 'How have all those exquisite adaptations of one part of the organic organisation to another part, and to the conditions of life, and of one distinct being to another being, been perfected?' (1859: 60). Unlike Paley, Darwin proposes that natural selection is the agent of this 'perfection of structure and coadaptation'.

Darwin set the tone for biologists who would later praise Paley for correctly recognising where evolutionary theorists should focus their explanatory efforts, and for understanding that mere chance alone could not possibly explain 'mechanism', at the same time as they note Paley's failure to identify the true nature of the process that led to these 'exquisite adaptations'.

Fisher, Maynard Smith, Dawkins

R. A. Fisher (in a 1948 letter to Nora Barlow) briefly praised Paley, while also hinting at his own recognition of how Paley had informed Darwin's work:

> The *Natural Theology* is full of material of interest to naturalists and displays Paley's wide interests in biological phenomena…[My] own guess is that he [Darwin] was quite considerably influenced by the *Natural Theology*... (Bennett 1983: 178).

Neo-Paleyanism is more clearly demonstrated in a series of remarks made by British adaptationists since the 1960s. John Maynard Smith avows Paleyan convictions about the primary job for the evolutionary theorist: 'The main task of any theory of evolution is to explain adaptive complexity, that is, to explain the same set of facts that Paley used as evidence of a creator' (1969: 82). Richard Dawkins is entirely explicit about his neo-Paleyanism, and I take the term from him:

> I suppose people like me might be labelled neo-Paleyists, or perhaps 'transformed Paleyists'. We concur with Paley that adaptive complexity demands a very special

> kind of explanation: either a Designer, as Paley taught, or something such as natural selection that does the job of a designer. (1998: 16)

Dawkins's (1986) decision to describe natural selection as a 'Blind Watchmaker' further underlines his debt to Paley. Natural selection, says Dawkins, has no foresight. In that sense it is blind, and that is why Dawkins subtitles his book, 'Why the Evidence of Evolution Reveals a Universe *without Design'* (emphasis added). Even so, natural selection still makes watches. Hence, for Dawkins, natural selection 'does the job of a designer', and thereby offers the same 'special kind of explanation' that true design offers to Paley.

Gardner and Grafen

Similar sentiments have been expressed more recently still. Andy Gardner has repeated the thought that Paley identifies the primary problem for evolutionary theory, while failing to give the correct solution: 'The problem of adaptation is to explain the apparent design of organisms. Darwin solved this problem with the theory of natural selection' (Gardner 2009: 861).[1] More recently, Gardner has augmented the basic Paleyan view, arguing that 'Darwinism is also a theory of the purpose of adaptation, i.e. the design objective of biological organisms. Darwin argued that…organisms will appear as if they are designed to maximise their survival and reproductive success…' (2017: 1). In other words, Darwin does

[1] Gardner does not mean to imply that Paley was the first person to identify this problem, merely that Paley stated it with particular clarity.

not merely show us how to explain the appearance of various traits that suggested good design to Paley, he also tells us what organisms' true design objective is.

We see the same combination of views from Grafen: 'Adaptation is the centre of biology, adaptation is design, and maximizing fitness is what organisms are designed for' (2007: 1248). Far from seeking to eliminate the notion of design from modern biology, Grafen agrees with Paley that adaptation is design. He moves on to give a very distinctive, and very precise, view about what the general purpose is of nature's designs.

3. Why Paley?

There are features of Paley's own approach which the neo-Paleyans I have drawn attention would not dream of endorsing. Take Paley's claim that, 'The air, the earth, the water, teem with delighted existence' (1802: 237-9). It is on the basis of the supposed superabundance of pleasure enjoyed by organisms, well beyond what is strictly necessary for survival, that he infers the goodness of the Deity. Darwinism in its various forms has often opposed this rosy view, suggesting instead to suggest that organic nature is characterised by struggle, in which far more are born than can possibly survive.

Paley sometimes asks question which the modern evolutionist would think absurd. For example, he wonders why animals have sensory organs at all, as opposed to simply perceiving the state of the world immediately (1802: 27). His answer is these organs have been given to animals in order to give us occasion to infer the creator's workmanship.

Neo-Paleyans can happily distance themselves from these sorts of remarks. They credit Paley simply with understanding that the world is characterised by good design, and also with understanding that this design demands a special kind of explanation that goes beyond chance. Their task is to articulate—much better than Paley—the true overall purpose of these designs.

4. The Eliminativist Challenge

Neo-Paleyans face more substantial worries about whether it makes sense to characterise the organic world using the language of design at all, once we have adopted a Darwinian outlook on life's organisation.

One reason why modern adaptationists reach for the language of 'good design' is to highlight the fact that organic traits are often highly tuned to enable various fitness-enhancing effects (see Lewens 2005). Even so, their conception of how this 'good design' is brought about is overtly opposed to Paley's image of a design process, and not merely because Paley thinks design involves intelligence. Paley claims that, 'every part of an animal or vegetable' has 'proceeded from a contriving mind' (1802: 47). Every part, 'is constructed with a view to its proper end and purpose' (ibid.) He insists on contrivance because it is implausible to think that:

> …the parts were not intended for the use, but that the use arose out of the parts. This distinction is intelligible. A cabinet-maker rubs his mahogany with fish-skin; yet it would be too much to assert that the skin of the fish was made rough and

> granulated on purpose for the polishing of wood, and the use of cabinet makers…But I think there is very little place for this in the works of nature. (1802: 40)

Paley insists that a design specification for the arrangement of the elements of organic nature must precede the construction of nature's mechanisms. It is not plausible, he insists, that nature primarily contains pre-existing elements turned to novel functions. To think this were true of nature in general would amount:

> …to such another stretch of assertion, as it would be to say, that all the implements of the cabinet-maker's workshop, as well as his fish skin, were substances accidentally configurated, which he had picked up, and converted to his use… (ibid.)

The very possibility that Paley dismisses—that organic nature consists in a series of pre-existing elements that have been opportunistically modified to suit new roles—is precisely the general mode of operation of natural selection as it acts on available variation.

Note, too, that while Paley thinks of 'design' in terms of the crafting of traits to fit their purposes, the logic of natural selection does not require that we think, in general, of the process of adaptation as the gradual shaping, or crafting, of organic traits to answer pre-existing environmental problems. Natural selection is—to an approximation that is good enough for our purposes in this section—a process whereby fitter variants proliferate in a population. Fitness might be augmented by adopting a slightly more favourable habitat, or by modifying the pre-existing ecological environment so that existing morphology is better suited to it (Lewontin 1978). No 'crafting' of form is required. Moreover, as the lineage and its environment are modified by selection, so the matter of what further opportunities are available for fitness-enhancement is liable to change. Thus, a general perspective on adaptation as an iterated process that results from differences in fitness between available variants fails to endorse anything closely analogous to Paley's notion of a design process; namely, a process whereby stable problems bring their solutions into existence, via the modification of organic morphology to meet those problems' demands.

Finally, the way in which Darwin conceives of selection and adaptation, compared with more modern conceptions, puts even greater distance between the *Origin*'s specific form of neo-Paleyanism, and today's general manner of conceiving of selection and adaptation.[2] Let us begin with a comparatively recent discussion of selection. Gillespie (1974) has shown that, so long as a population is finite, selection can favour low variance in offspring number. To see why this is the case, consider this simplified example, adapted from Sober (2001). Imagine a population that contains two types. The As always have 2 offspring. The Bs either have 1 or 3, with equal probability. We can show that in spite of the expected number of offspring of each individual—whether A or B—being identical, we should nonetheless bet on A increasing its frequency in the population over B. In generation one, imagine that we have 2 As and 2 Bs. In generation two we can be sure there will be 4 As. What of the Bs? There is a 25% chance of 2Bs, a 50% chance of 4 Bs, and 25% chance of 6 Bs. Crucially, it is because the frequency in the population of As is determined in part by population size, which in turn is determined by the number of Bs, that we should expect A to increase its frequency in generation 2 (to roughly 52%).

[2] See also Lewens (2010).

For Gillespie this is an instance of natural selection favouring lower variance. That is because the genotype that bestows higher variance has a lower fitness 'as measured by its mean frequency in the next generation' (1977: 1012). The fact that the genotype associated with lower variance reliably increases in frequency indicates, for Gillespie, that we have an instance of natural selection favouring that genotype. What is more, the opening sentences of Gillespie's original (1974) treatment of selection for within-generation variance in offspring number motivate the problem by asking us to consider whether a bird should have one large clutch in a breeding season, or many smaller ones. Should the bird, as he puts it, put all its eggs in one basket, or distribute them among several? Here we have a pair of strategies, with better and worse options. Our stylistic presentation of the case of selection favouring lower variance does not lend itself immediately to an interpretation in terms of good design in the sense of suitableness to some ecological demand. Even so, it is possible to regard the very fact that the As are likely to make a greater proportionate contribution to the next generation as sufficient justification for thinking of them as better 'designed' than the Bs. Grafen (1999) goes further, attempting to find a single very general notion of fitness, such that in all circumstances, including the kind of cases discussed by Gillespie, that is the quantity favoured by selection. Hence Grafen conceives of his project as the articulation of a general *criterion* of design.[3]

---

[3] The term 'criterion' is Grafen's own: 'The main purpose of [Fisher's fundamental] theorem is to find a quantitative measure of the effect of natural selection in a Mendelian system, thus founding Darwinism on Mendelism and *identifying the design criterion for biological*

It is important, nonetheless, to note how far we have moved here from Darwin's understanding of natural selection. Grafen's efforts to preserve a link between selection and design are underpinned by the assumption that quality of design can be measured in terms of some quantitative contribution to the composition of the next generation. The difference between this view and Darwin's own is indicated by Darwin's tendency to draw a fairly strong distinction between natural selection and sexual selection. For more modern theorists sexual selection is more usually understood as a variety of natural selection.

For Darwin, natural selection 'works solely by and for the good of each being' (1859: 489). It acts on variations that are 'useful to each being's welfare' (ibid.: 127). These references to variations that improve *welfare*, or a being's *good*, are not stylistic accidents of Darwin's, nor are they simply metaphorical ways of talking about an organism's reproductive success. Darwin refuses to use the term 'natural selection' to name a process that generally favours traits that make reliable contributions to an organism's reproductive success. Darwin's refusal is exemplified by his insistence that sexual selection, unlike natural selection, can promote traits that depend 'not on a struggle for existence, but on a struggle between the males for possession of the females' (1859: 88). For Darwin, while natural selection acts to favour variations that promote 'each being's welfare', sexual selection can instead undermine this tendency (1859: 127). This is in spite of the fact that both sexual selection and natural selection favour traits which assist in an organism's contribution to the next

---

adaptation, embodied in Fisher's ingenious definition of fitness.' (Grafen 2015: 1, emphasis added.)

generation, hence which meet the criterion of good design as Grafen would understand it. Darwin thought of natural selection in this comparatively narrow way because he was interested in giving a general argument for how traits that promote individual welfare can be built over time. This was his goal because he aimed to account for the very same beneficial traits that Paley was interested in (Lewens 2018).

The arguments in this section—all of which point to significant differences between the manner in which Paley conceived of design as process and product, and the general modern understanding of selection and adaptation—help to diagnose why the neo-Paleyan approach has attracted regular attack. They also point the way to rebuttals of those attacks. There are inevitably similarities and differences between any two domains, and those who attack the design approach tend to focus on the differences, while those who defend the approach instead stress the similarities. Ollason, for example, published a series of articles through the 80s arguing that all notions of design ought to be, but had not yet been, eliminated from biology (e.g. Ollason 1987). More recently, Reiss (2009) has argued for a similar elimination of 'design'. He is concerned (among other things) to stress the difference between a conception of design according to which a stable, pre-existing problem guides the crafting of its own solution, and the quite different manner in which selection can lead to the proliferation of whichever available variant happens to promote fitness in a set of local circumstances. In support of Reiss, we have seen that it is not true that natural selection always works by crafting the form of an organic trait to fit a stable pre-existing problem.

Eliminativism also receives a boost when we ask whether biologists need to frame their technical presentations of evolutionary work in the language of design. A physicist might point out that there are features of the cosmos that were previously thought to be evidence for the existence of a designer—the orderly workings of the solar system, for example—which physics is now able to explain without invoking any overseeing intelligence. Even so, that physicist would be very unlikely to advocate that we continue to think of the universe in terms of good design, or that a technical theory of design is still required in physics. While there is nothing objectionable about a biologist claiming, with pride, that she can explain phenomena once thought indicative of intelligent oversight, why not simply remove ongoing references to design in modern technical theorising itself, perhaps replacing them with different notions of selection, fitness and so forth?

As we will shortly see, the disanalogies between Paleyan notions of design and modern notions of natural selection and fitness do not show that the neo-Paleyan approach is impermissible in all its forms. The goal of pinning down a quantity most generally favoured by natural selection, hence a quantity which, for that reason alone, deserves to be understood as a 'design criterion' of evolution, remains in play given all of the objections we have seen so far. It is this aspect of the design-based approach that Gardner and Grafen, our most recent neo-Paleyans, have accentuated. The next section, a preliminary to an examination of this appeal to a design 'criterion', examines their responses to the eliminativist challenge.

5. Gardner on Circularity and Tautology

Theorists have sometimes argued that adaptation should be defined in terms of fitness and selection. Elliott Sober, for example, asserts that: ‘A is an adaptation for task T in population P if and only if A became prevalent in P because there was selection for A, where the selective advantage of A was due to the fact that A helped perform task T’ (Sober 1984: 208). Sober offers the prospect of understanding adaptation in terms of selection, and selection in terms of fitness. His remarks thereby raise the issue of whether we might dispense with the notion of design altogether, in favour of a different family of technical notions.

In contrast, we have already seen a comment from Grafen suggesting that we should define adaptation in terms of design, not selection: ‘…adaptation is design, and maximizing fitness is what organisms are designed for’ (2007: 1248). Gardner offers two related arguments in support of Grafen’s approach.

First, Gardner suggests that defining adaptation in terms of natural selection confuses the issue of what adaptation *is*, with the issue of how we should go about explaining it:

> …many researchers have defined adaptation as *any* response to natural selection. However, this definition confuses the original empirical problem (adaptation) with its solution (natural selection). (Gardner 2009: 861)

This has resonances with an earlier argument from D. C. Fisher (1985). He argued that if we define ‘adaptation’ as any trait that is a product of natural selection, then we immediately render Darwin’s claim that natural selection explains adaptation a trivial consequence of that definition. It is partly for these reasons that Gardner praises Paley: 'Paley explicitly framed the problem of adaptation as the need to explain the empirical fact of apparent organismal design' (ibid.). We should understand the central problem of natural history in terms of the explanation of 'apparent organismal design', says Gardner, because this preserves the empirical significance of Darwin's claim that it is natural selection that explains this phenomenon.

If we define 'adaptation' as 'structure explained by natural selection', then it is indeed empty to assert in general that natural selection is what explains adaptation. This does not render specific claims vacuous, such as the claim that natural selection is what explains how eyes, organs of echolocation etc. come to be. Moreover, this selection-based definition of adaptation allows us to retain the claim that, as a matter of fact, natural selection is the correct explanation for many puzzling phenomena that previously gave encouragement to natural theologians, which they previously conceived of in terms of design, and which we now refuse to understand in the language of design. We can coherently refuse to define 'adaptation' in terms of 'design', without rendering Darwin's most important insights trivial (Lewens 2007).

Gardner has a second argument for thinking of Darwinism in terms of design, which begins with his consideration of the 'tautology problem'. The supposed worry about tautology begins with the (naïve) notion that the principle of natural selection tells us that the fittest survive. This formulation is a tautology if we understand 'fitness' in terms of actual survival. Switching to a propensity definition of fitness—understood not in terms of actual number of offspring produced, but in terms of a chancy disposition to produce offspring—does not solve the tautology problem. It is, after all, false that the fittest in this sense invariably out-reproduce the less fit. In small populations, drift can intervene. If we revise our naïve understanding of the principle of natural selection, and replace it with the claim that, in large populations in the long run, the fittest are the most likely to proliferate, then we still have a tautology. Our new version of the principle expresses the general logical relationship between propensities and long-run frequencies (see Sober 1984).

Gardner responds to the tautology problem:

> The apparent tautology vanishes when one understands that Darwinism is really a theory of organismal design. It is possible to imagine worlds where organisms are designed to boil water or write poetry. The empirical value of Darwinism owes to it identifying that, in this world, organisms are designed to achieve reproductive success. (2009: 861-2)

Some evolutionary psychologists have suggested that just as birds have been selected to produce a melodious song, humans really have been designed to write poetry, because (in their view) this contributes to reproductive fitness via its impact on attracting mates (e.g. Miller 2000). Evolution by natural selection can, in principle, favour any trait—artistic endeavour, even the boiling of water—so long as it is somehow linked to reproductive output. Gardner's point, of course, is not to undermine any specific claim that might be thrown up by evolutionary psychology about the adaptive value of poetry. Instead he suggests that we can answer the tautology problem by construing Darwin's theory, and subsequent refinements of the theory of natural selection, as accounts that posit wholly general objectives for all organismal design.

In response to this claim, we should note that the tautology problem can be answered without recourse to a notion of design. Previous theorists (e.g. Sober 1984) have noted that even if the principle of natural selection expresses a tautology, there remains plenty of empirical content in the specific claim that (for example) the melanic form of *Biston betularia* proliferated in the woods of Derbyshire because it had an advantage in avoiding predation from birds. So a general account of organismal design is not required to answer the tautology problem. Nevertheless, there may be other grounds for its justification, to which we now turn.

6. Adaptationism, Optimality and Design

Gardner has suggested that a fruitful form of adaptationist thinking requires us to be able to articulate a general design criterion for evolution. He identifies what he calls *strong adaptationism* as 'a caricature of Darwinism in which organisms are regarded as entirely optimal…' (2017: 2). The status of this view as a caricature does not mean that Gardner writes it off. His view is that this version of adaptationism is false if we think of it as a hypothesis, but it can nonetheless be important when used as 'an investigative tool'. As Gardner puts it, '…by an iterative process of model adjustment, testable prediction and empirical test, the optimization approach provides an investigative tool by which scientists learn how the biological world works' (ibid.).

This broad approach to optimality thinking has distinguished defenders: Oster and Wilson remarked long ago that, '. . . optimization models are a method for organizing empirical evidence, making educated guesses as to how evolution might have proceeded, and suggesting avenues for further empirical research' (Oster and Wilson 1984: 273). Gardner gives a specifically neo-Paleyan gloss on how this heuristic form of optimality thinking works: 'Without knowing what organisms are designed to do, it would be impossible to decide which of a range of possible phenotypes represents the optimum' (2017: 2).

Gardner seems to be suggesting that we have a straight choice between a neo-Paleyan version of adaptationism, which grounds our thinking about nature in a general account of 'what organisms are designed to do', versus a woolly and unproductive form of anti-adaptationism. As he puts it: 'The maladaptation view is strictly correct but also completely useless if it does not yield specific, testable predictions' (2017: 2).

We should follow Gardner in defending the benefit of models that yield 'specific, testable predictions'. Sober's (1984) discussion of what he calls 'censored forces models'—i.e. models which describe what would happen given some edited subset of the forces suspected to operate in an evolutionary population—reinforces this position. We can generate valuable hypotheses about the actual processes affecting real populations by seeing how close our simplified model-based predictions are to the actual outcomes seen in a population, and then revising our assessment of the forces in play if we note significant differences between the two. But it does not follow from this account of the benefits of models, quantification and simplification that we must approach nature in a manner that begins by asking what populations would be like, if they were subject only to selection operating according to some general design criterion.

First, we can build a model that shows how a population would evolve, if it were subject only to drift (Lewens 2009). If the population departs from what we expect from a quantitative model of drift acting alone, then we can conjecture that other forces—selection, migration etc.—must also have been at work. Our opening question need not be the adaptationist one, 'What would nature be like if design were optimal?', even if the end-point of our investigation is that we conclude that selection has acted in important ways.

Second, the very general heuristic strategy of conjecturing some outcome based on a quantitatively specified set of evolutionary forces, and then revising our assessment of actuality if our model-based predictions depart from reality, does not rely on our having a wholly general theory of what organisms are designed to do. Instead, we can conjecture, in a quantitatively rigorous way, expected evolutionary outcomes in specific circumstances based on local assumptions about inheritance, group structure, reproductive output and so forth. We can build a series of different models, which aim to evaluate different local scenarios. Here, too, we will develop what Gardner values—namely 'specific, testable predictions'—but with no need for anything like a general design principle to inform our efforts.

I suggest, then, that Allen et al. are strictly correct when they attack the neo-Paleyan position with the insistence that, '…no universal maximands or design principles are needed to understand the evolution of social behaviour'. Universal design principles are not needed for understanding of any aspect of evolution. Instead, as they put it, we need 'modelling assumptions. These assumptions can be highly specific, applying only to particular biological situations, or broad, applying to a wide range of scenarios' (2013: 20138). But even if it is not strictly necessary for studying evolution, if a criterion of 'what organisms are designed to do' could be articulated and shown to hold with generality, then that criterion would be of great value.

7. The Fundamental Theorem

Gardner believes there is such a general criterion:

> Explaining the purpose of adaptation is a central achievement of Darwinism. Being able to predict what it is that organisms are striving to achieve not only sets Darwinism apart from intelligent design and other forms of mysticism, but also sets the hugely successful adaptationist research programme apart from scientifically sterile anti-adaptationist thinking within evolutionary biology. (2017: 6)

Gardner turns to Fisher at this point. Here he is following earlier remarks from Grafen, who tells us that '[Fisher's] fundamental theorem tells us what it is that the design-creating capacity of evolution regards as good design' (Grafen 2003: 326). What, then, is the Fundamental Theorem?

Fisher's best-known formulation of the Theorem is this one:

> The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time. (Fisher 1930: 35)

There are some strange turns of phrase in this formulation—the increase in fitness of *an* organism, *an* organism's variance in fitness—which (as other commentators have pointed out) we are best off ignoring (see Okasha 2008). Fisher's own later formulations of the theorem all suggest any worries we have about these terms merely reflect odd stylistic choices on his part. His 1941 formulation, for example, is explicit in speaking of populations: 'The increase in average fitness of the population ascribable to a change in gene frequency…is equal to the genetic variance in fitness' (Fisher 1941: 377).

This is not the place to offer anything like a full interpretation of the Fundamental Theorem (see Price 1972, Edwards 1994, 2016, Plutynski 2006, Okasha 2008, Grafen 2015, 2018, Ewens and Lessard 2016 for important contributions). Instead, I shall rely on some broad points of consensus. Variance is a statistic that can never be negative. But the Fundamental Theorem does not tell us that the mean fitness of a population that is affected solely by natural selection must always go up. It is just as well that it does not tell us this, because there are uncontroversial cases where a population is affected solely by natural selection and yet its mean fitness goes down. This is easiest to see when we focus—as Gardner (2017) also does—on frequency-dependent selection. Consider a very simple example, where we imagine a population containing two behavioural types—the bellicose, who like to fight for resources, and the pacific, who run away from contests. We can suppose that the bellicose are fitter than the pacific. As a result, they increase in frequency. But they consequently encounter each other more often, they become constantly involved in gruelling fights, and population mean fitness declines.

The standard modern approach, which draws on Price's seminal analysis (Price 1972), argues that we should not interpret the Fundamental Theorem in such a way that these sorts of cases demonstrate its falsehood. Okasha, for example, argues that the Fundamental Theorem concerns, '…the partial rate of change which results from natural selection altering gene frequencies in a population, *in a constant environment*' (Okasha 2008, emphasis in original). On this view, the example just outlined is no counter-example because the environment is not fixed: the frequency of bellicosity—a quantity that increases in each generation—is itself an environmental feature. What is more, Fisher's conception of the 'environment' is, as Price pointed out, exceptionally broad. It includes every factor that disrupts constant average effects of allele substitutions on fitness. This means that under the heading of 'environment' we can include not just the concentration of behavioural types in a social environment, or facts relating to population density, but also factors that are internal to organisms such as genetic dominance and epistasis.

When Okasha writes that the Theorem only applies in 'constant' environments, this may create the impression that the Theorem never applies to real populations, for in reality the environment is always changing. This would be a misreading of the Fundamental Theorem and, I suspect, of Okasha's intention in pointing to environmental constancy. Grafen and Gardner more typically write of the Fundamental Theorem as concerning the effects of selection at any instant in time. Grafen puts it like this: 'the measurement of natural selection is made in the circumstances of the moment, and does not look ahead to changes in the population constitution, even when those changes are brought about by natural selection itself' (2018: 181). So, again, the example of frequency dependent selection above is no counter-example to the Theorem. The fittest trait is favoured just as the Theorem tells us, but as a consequence of this the population itself is altered in such a way that mean fitness is subsequently reduced.

Price summarised the Fundamental Theorem thus:

> What Fisher's theorem tells us is that natural selection…at all times acts to increase the fitness of a species to live under the conditions that existed an instant earlier. But since the standard of 'fitness' changes from instant to instant, this constant improving tendency of natural selection does not necessarily get anywhere in terms of increasing 'fitness' as measured by any fixed standard. (1972: 131)

He further suggested that in reality mean fitness '…is about as likely to decrease under natural selection as to increase' (ibid.) Note, crucially, that Price's comment relates to the anticipated effects of natural selection when unimpeded by other forces. The simple case of frequency dependent selection outlined above reminds us that even when selection is the only evolutionary force at work, we should still have no general expectation that fitness will increase over time. As Gardner points out, this is a case where the primary effect of selection is to increase bellicosity, but what he calls a 'secondary consequence', also of selection, is that this ‘deterioration in the social environment results in a net decrease in average fitness’ (2017: 2).[4]

8. Design and the Fundamental Theorem

Fisher was well aware of what could not be concluded on the basis of the Fundamental Theorem. He understood that selection’s effect needed to be weighed against tendencies that undermine increases in fitness:

> Against the action of Natural Selection in constantly increasing the fitness of every organism…is to be set off the very considerable item of the deterioration of its inorganic and organic environment. (1930: 45)

---

[4] Grafen (2018: 185) credits Dawkins with, ‘The whole understanding that natural selection operates at the level of individuals, and that mean fitness can go down as a result of natural selection operating in a frequency-dependent environment.’ Let me be clear, then, that it is not my aim in this section to imply that what Grafen and Gardner say about the Fundamental Theorem is in any way challenged by the case of frequency-dependent selection. Both theorists explicitly note the well-known points made here. Jean-Baptiste Grodwohl has been especially helpful in suggesting improvements to this section.

This reference to selection's action 'in constantly increasing the fitness of every organism' may explain Price's own characterisation of natural selection as the only evolutionary force with a 'constant improving tendency' (1972: 131). In saying that selection is the only force with this tendency, we should not understand this to mean that its net effect will be to increase fitness so long as other forces do not intervene. We have already seen that this is not the case, because we have seen that it can result in fitness going down. Instead, it is the only force with an improving tendency in the sense that there is no *other* evolutionary force that has a *general* tendency to increase fitness at any instant in time. When Grafen (2003: 326) writes that, 'Fisher thought that his fundamental theorem isolated what we might call the adaptive engine of Darwinian natural selection', he is not making a claim about selection's likely long-run effects on adaptive improvement. Instead he is pinning down how selection differs from other forces.

Grafen goes on to say that 'Fisher believed that [the] partial change [due to selection] was the only aspect of the changes in a population's genetic constitution that was progressive, that could create design.' When Fisher defended the overall status of the Fundamental Theorem in 1953 (see Edwards 2016), his remarks are consistent with Grafen's claim about what makes selection unique. They also acknowledge the countervailing effects on actual population outcomes of what Fisher here calls 'a deteriorating climate':

> In this way the principle of natural selection is put upon an exact quantitative basis, and the course of spontaneous population change is related to its state at any instant. The improvement in relation to a fixed external environment is always positive… Of course, for any particular species, it does not follow that it is gaining on

> its competitors and enemies, or even that its increasing efficiency is enough to meet the needs of a deteriorating climate. Natural selection does not imply that every species is progressively better off, but that all species tend by all possible changes of gene ratio to improve their average condition. (Fisher 1953)

What Fisher does not stress here, but which is important for evaluating some aspects of the neo-Paleyan approach to evolution, is that the very forms of environmental deterioration that undermine the improving tendencies of natural selection can be downstream effects of earlier episodes of selection itself.[5]

Gardner is well aware of this, as is Grafen. We have seen that they both make the same point explicitly (see Gardner 2017, Grafen 2003, and the passages already quoted from Grafen 2018). When Grafen (2003: 327) suggests that the importance of the Fundamental Theorem lies in the possibility that it might ultimately justify biologists' 'informal sense that natural selection leads to organisms maximising their fitness', he is well aware that such a maximising tendency must be understood in a manner that is compatible with selection decreasing the mean fitness of a population over time. Even so, some readers might be misled by some of Gardner's remarks on the link between Fisher and the neo-Paleyan project. Gardner claims that:

---

[5] Fisher does, however, make such a remark in his (1941: 53) when he notes that, 'A change in the proportion of any pair of genes itself constitutes a change in the environment in which individuals of the species find themselves.' I thank an anonymous referee for drawing my attention to this comment.

> [Fisher] showed that natural selection always acts to increase the mean fitness of the population, in what he termed the 'fundamental theorem of natural selection' (FTNS). Fisher interpreted this as proof that organisms will appear increasingly designed as if to maximise their Darwinian fitness. (2009: 862)

Gardner's comments elsewhere remind us we must understand 'always acts to increase', in a manner compatible with the overall result of selection over time being the *reduction* of a population's mean fitness. Turning to Gardner's interpretation of Fisher, we have seen that while Fisher does talk about that 'action of natural selection in constantly increasing the fitness of every organism', and while he also uses the language of 'improvement', he does not explicitly use the language of design. Fisher also reminds us of the various factors that can offset selection's improving tendency, thereby raising the question of whether organisms really will appear 'increasingly designed' to maximise their fitness. But what is clear—and uncontroversial, given how we must interpret the Fundamental Theorem in order to render it compatible with the simple case of frequency-dependent selection noted by Gardner—is that the Fundamental Theorem itself does not prove that design quality will generally go up, even when natural selection acts without the influence of other evolutionary forces.

9. Justifying the Design Heuristic

We have already noted Grafen's (2003: 327) argument that the Fundamental Theorem is important today because: 'most take for granted an informal sense that natural selection leads to organisms maximizing their fitness, but they do not ask how that sense can be justified.' Later (2014a: 156) he describes this in more explicitly neo-Paleyan terms as an approach 'based on the expectation of finding good design in nature: this stretches back at least to natural theology in the eighteenth century, and was invigorated and reinforced as a scientific approach by Darwin (1859) and later Fisher (1930)'. Grafen has aimed, both through his work on Fisher, and through his 'Formal Darwinism' project, to provide a theoretical justification for those who think 'it is productive to regard organisms as well-designed in relation to their environment' (Grafen 2014a: 167; see also Okasha and Paternotte 2014).

There is good evidence that some kind of presupposition about good design characterises much research in fields such as behavioural ecology. Davies, Krebs and West typify the approach:

> Individuals should be designed by natural selection to maximise their fitness. This idea can be used as a basis to formulate optimality models… (2012: 81)

Birch (2016) gives far more evidence for the broad acceptance of this idea. Note that the term 'design' features explicitly in these comments from behavioural ecologists, and not merely the notion that we should expect selection always to promote fitness. This is in line with a long-standing neo-Paleyan tendency to approach the analysis of organisms in a manner that parallels how we might analyse designed artefacts. As Krebs and Davies put it:

> Visitors from another planet would find it easier to discover how an artificial object, such as a car, works if they first knew what it was for. In the same way, physiologists are better able to analyse the mechanisms underlying behaviour once they appreciate the selective pressures which have influenced its function. (Krebs and Davies 1997: 15)

We have seen that, for Grafen and Gardner, a proper understanding of the Fundamental Theorem allows us to appreciate which quantity is most generally favoured by selection. In this sense they also aim to show selection's 'improving tendency' from one moment to the next. We have also seen that Grafen and Gardner do not equate this with the justification of an expectation—even in the absence of other forces—that selection will increase mean fitness over time. Now consider Grafen's comments on his 'Formal Darwinism' project:

> A major point in my presentation of the formal darwinism project…is that behavioural ecologists use fitness-maximisation ideas routinely, should welcome a justification of this practice, and may sometimes be in need of guidance about precisely what good design means in an organism… (Grafen 2014b: 286)

To give an account of 'what good design means' in an organism is to give a general account of a quantity that is most generally favoured by selection (see Okasha and Martens 2016 for concerns about recent efforts to specify this quantity with precision). This is the job of supplying a general 'criterion' of good design, which, as we have seen, Grafen and Gardner argue that the Fundamental Theorem can help us with. The question is to what extent the justification of such a criterion also amounts to a justification of behavioural ecologists' practice of using the notion of design. The problem is that behavioural ecologists do not merely assume that when selection acts, it always acts to favour some form of fitness that can be stated with generality. Hence they are not merely investigators in search of a general design criterion. In addition, it appears that some of them, at least, also assume that we can fruitfully approach the organic world in a way that assumes that reasonably good design solutions emerge over time in response to selective pressures. That seems to be the way in which Krebs and Davies, for example, recommend that we approach the natural world. Their suggestion is not merely that selection always promotes fitness at any instant in time. That would be compatible with selection always have medium and long-term effects on environments that undermine the production of anything we might characterise as 'design'. Instead, they suggest we can fruitfully see the environment as posing specific problems which are then solved in the lineages exposed to those environments. This approach to design is most explicit in these widely-cited methodological comments from evolutionary psychologists Barkow, Cosmides and Tooby:

> By understanding the selection processes that our hominid ancestors faced – by understanding what kind of adaptive problems they had to solve – one should be able to gain some insight into the design of the information-processing mechanisms that evolved to solve these problems (Cosmides, Tooby and Barkow 1992: 9).

Grafen and Gardner's conclusions regarding the criterion of design have considerable potential value for behavioural ecologists. Their aim is to use the Fundamental Theorem in ways that tell the field researcher that there is only one quantity that selection always favours, hence any adaptation must contribute to fitness as understood in the manner they recommend. Even so, in focusing on justifying the criterion of good design in this way, Grafen and Gardner do not thereby justify any stronger expectation of finding good design in nature; nor do they licence an approach that sees organic traits as solutions to environmental problems, or as products designed in response to selection pressures (see also Birch 2014, Okasha 2018).

The question of whether a selection process produces something that suggests good design depends not merely on whether selection consistently favours a single quantity, but also on further facts about which the Fundamental Theorem itself is silent.[6] To give just a few examples, there must be enough consistency over time in the challenges set by the environment for payoffs to consistently favour cumulative trajectories of improvement. There must also be a suitably rich supply of variation to enable these trajectories to be followed. Lewontin further suggests, in his discussion of what he calls 'quasi-independence',

[6] In writing that the Fundamental Theorem is silent on these further questions I do not mean to imply that Fisher is silent on all of them, or even that he is silent on these further matters in *The Genetical Theory of Natural Selection*. For example, the section immediately following Fisher's presentation of the Fundamental Theorem contains an important further geometrical argument, relevant to the questions treated here: 'Adaptation, in the sense of conformity in many particulars between two complex entities, may be shown, by making use of the geometrical properties of space of many dimensions, to imply a statistical situation in which the probability, of a change of given magnitude effecting an improvement, decreases from its limiting value of one half, as the magnitude of the change is increased' (1930: 46).

that if developmental organisation is too highly integrated, then this may also disrupt the pursuit of cumulative trajectories (Lewontin 1978). The widely discussed problem here is that if, for example, mutations that might potentially produce improvements to a hypothetical proto-eye also tend to have additional effects on a host of other traits in the organism, then likely adverse fitness effects with respect to these other traits can be expected to outweigh any positive fitness effect with respect to vision. In brief, then, there is a significant gulf that separates our ability to secure a claim about what we have been thinking of as a general design 'criterion'—in the sense of a quantity that selection generally promotes—and justifying 'the expectation of finding good design in nature'. Since the Fundamental Theorem says nothing about the additional factors that are required to bridge this gulf, it cannot by itself vindicate the version of neo-Paleyanism that seems to be indicated by the comments of Krebs and Davies, and Cosmides and Tooby.

10. Explaining the Design Heuristic

We can make headway by switching from trying to justify the design heuristic, to trying to explain the design heuristic (see Birch 2018, Okasha 2018). Behavioural ecologists who work in neo-Paleyan mode have, as a matter of fact, found it fruitful to approach nature through the lens of design. They have found it useful, as we have seen, to build models that assume that organic traits are well-designed solutions to specific environmental problems. The successful track record of this approach is sufficient to justify its continued use, and because of this a formal theoretical justification is not strictly necessary for behavioural ecologists to feel comfortable in their practice.[7] We now see that what we have is an explanatory question: what must the biological world be like, and what must investigators be like, such that thinking in terms of good design is empirically fruitful and psychologically attractive (Lewens 2004, Birch 2014, 2016)?

A very general principle such as the Fundamental Theorem evidently cannot give us a complete answer to these questions.[8] We must turn to facts about (for example) the organisation of organic developmental systems, the abundance of variation and so forth if we are to understand why it is that as a matter of fact the design heuristic has payoff. Ultimately, this also explains why the role that the Fundamental Theorem can potentially play in the neo-Paleyan project—namely as a way to clarify the quantity that selection systematically promotes at any instant—is described as justifying a *design* criterion, as opposed to a criterion with no link at all to design. It is possible to imagine a biotically impoverished world where selection systematically promotes fitness from one instant to the next just as the Fundamental Theorem describes, and yet (for example) variation is so highly constrained that nothing that puts us in mind of 'good design' emerges at all. Under such circumstances we *might* conceivably think of the Theorem as justifying a design criterion

[7] Gardner makes a very similar point: '…the strong adaptationist approach must ultimately be judged upon its ability to facilitate empirical testing of basic theory, and provide new biological insights, and not upon the biological face value of its underlying axioms' (2009: 863-4).

[8] Let me stress, once again, that many of Grafen and Gardner's comments quoted in this essay indicate that they may well agree with this assessment.

which, for extraneous reasons, never succeeds in producing good design. Alternatively—and, I suggest, more plausibly—while we might still characterise the Theorem as pinpointing a quantity generally favoured by selection, we would be unlikely to think of this as delivering a *design* criterion at all (see also Birch 2014).

11. Vindicating Darwin's Argument

Gardner suggests a final role for the Fundamental Theorem. It gives a formal grounding to Darwin's neo-Paleyan argument in the *Origin*. Gardner believes that Darwin tells us that: 'over successive generations, organisms should appear increasingly well designed for the purpose of achieving reproductive success…' (2009: 861). There is an ambiguity in this formulation. Gardner may simply be attributing to Darwin the view that there is no quantity *other than* reproductive success that is generally favoured by evolutionary processes. This is roughly the view Gardner defends in his discussion of the FTNS. Alternatively, Gardner may be attributing a much stronger view to Darwin, namely that it is reasonable to expect that designs of increasingly high quality will emerge over successive generations. We have already seen that these two formulations are not equivalent, and that the second formulation is not justified by the FTNS.

Darwin may well endorse the stronger view:

> It may be said that natural selection is daily and hourly scrutinizing, throughout the world, every variation, even the slightest; rejecting that which is bad, preserving and adding up all that is good; silently and insensibly working, wherever and whenever

> opportunity offers, at the improvement of each organic being in relation to its organic and inorganic conditions of life. (1859: 84)

If we read Darwin's remarks in a full-blooded way, as asserting that natural selection can be reliably expected to 'add up all that is good' for individual organisms, then those remarks are not justified by the Fundamental Theorem, for reasons already outlined (see also Birch 2014).

Fortunately, Darwin's achievement in undermining his special creationist opponents does not rely on such strong interpretations of natural selection's improving tendencies. Darwin needed to show that when we find traits that previously suggested intelligent design, they could be explained as the results of iterated cycles of natural selection (see Birch 2014, 2016, 2017, Okasha 2018). This can be true even if there is no general expectation that natural selection reliably produces traits that appear well designed.

The well-trodden example of syphilis and paresis supports this point. Famously, very few individuals who have late-stage syphilis—the proportion is around 10%—go on to get paresis (Scriven 1959). And yet, since people get paresis only as a later consequence of syphilis, it is usually taken to be uncontroversial that we can explain why an individual has paresis by reference to their syphilis. Syphilis can explain paresis, even though syphilis does not in general lead to paresis. Similarly, natural selection can explain the appearance of adaptations such as the eye, even in the absence of any general tendency for selection to produce traits that give the appearance of good design.

12. Conclusions

This paper has drawn attention to a very long-standing 'Neo-Paleyan' tradition in British evolutionary theorising, which began with Darwin and continues to the present day. We have seen a number of significant disanalogies between Paleyan conceptions of design and modern conceptions of adaptation and selection, which help to explain why the neo-Paleyan programme is sometimes treated with hostility. These general disanalogies do not suffice to dismiss the most interesting forms of recent neo-Paleyanism, which tend to draw on theoretical principles such as Fisher's Fundamental Theorem to ground a general approach to what we have called (following Grafen) the 'criterion' of evolutionary design. We have seen that we must sharply distinguish efforts to justify this design criterion from efforts to show that selection reliably produces good design. We saw that the design heuristic can be justified by its success, and that these successes raise an important set of explanatory questions about why that heuristic has been successful which must draw on resources that go well beyond the Fundamental Theorem itself. Finally, we saw that Darwin's argument for selection as an explanation of adaptation does not require that selection in general produces traits that suggest good design to human observers.

References

Allen, B., M. Nowak and E. O. Wilson (2013) 'Limitations of Inclusive Fitness' *PNAS* 110: 20135-20139.

Bennett, J. (1983) *Natural Selection, Heredity and Eugenics, Including Selected Correspondence of R. A. Fisher with Leonard Darwin and Others*. Oxford: Clarendon Press.

Birch, J. (2014) 'Has Grafen Formalised Darwin?' *Biology and Philosophy* 29: 175-180.

Birch, J. (2016) 'Natural Selection and the Maximisation of Fitness' *Biological Reviews* 91: 712-727.

Birch, J. (2017) 'The Inclusive Fitness Controversy: Finding a Way Forward' *Royal Society Open Science* 4: 170335.

Birch, J. (2018) 'Fitness Maximisation' in R. Joyce (ed.) *The Routledge Handbook of Evolution and Philosophy*. London: Routledge.

Birch, J. (forthcoming) 'Inclusive Fitness as a Criterion for Improvement' *Studies in History and Philosophy of Biological and Biomedical Sciences*.

Cosmides, L., J. Tooby and J. Barkow (1992) 'Introduction: Evolutionary Psychology and Conceptual Integration' in J. Barkow, L. Cosimdes and J. Tooby (eds.) *The Adapted Mind*. Cambridge, MA: MIT Press.

Darwin, C. (1859) *On the Origin of Species*. London: John Murray.

Darwin, C. (2002) *Autobiographies*. M. Neve and S. Messenger, eds. London: Penguin.

Davies, N. B., Krebs, J. R. & West, S. A. (2012). *An Introduction to Behavioural Ecology*. Wiley-Blackwell, Hoboken.

Dawkins, R. (1986) *The Blind Watchmaker.* London: Penguin, 1986.

Dawkins, R. (1998) 'Universal Darwinism' in *The Philosophy of Biology*, D. Hull and M. Ruse (eds). Oxford: Oxford University Press.

Edwards, A. W. F. (1994) 'The Fundamental Theorem of Natural Selection', *Biological Reviews* 69: 443–74.

Edwards, A. W. F. (2016) ‘Analysing Nature’s Experiment: Fisher’s Inductive Theorem of Natural Selection’ *Theoretical Population Biology* 109: 1-5.

Ewens, W.J. and S. Lessard (2015) ‘On the interpretation and relevance of the Fundamental Theorem of Natural Selection’ *Theoretical Population Biology* 104: 59–67.

Fisher, D.C. (1985) ‘Evolutionary Morphology: Beyond the Analogous, the Anecdotal, and the Ad Hoc‘ *Paleobiology* 11: 120–138.

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

Fisher, R. A. (1941) ‘Average Excess and Average Effect of a Gene Substitution’, *Annals of Eugenics* 11: 53–63.

Fisher, R.A. (1953) ‘Population genetics’ *Proceedings of the Royal Society of London Series B* 141: 510–523.

Gardner, A. (2009) ‘Adaptation as Organism Design’ *Biology Letters* 5: 861-864

Gardner, A. (2017) ‘The Purpose of Adaptation’ *Interface Focus* 7: 20170005.

Gardner, A. and J. Welch (2011) ‘A formal theory of the selfish gene’ *Journal of Evolutionary Biology* 24: 1801-1813.

Gillespie, J. (1974) ‘Natural Selection for Within-Generation Variance in Offspring Number’ *Genetics* 76: 601 - 606.

Gillespie, J. (1977) ‘Natural Selection for Variances in Offspring Numbers: A New Evolutionary Principle’ *American Naturalist* 111: 1010-1014.

Gillespie, N. (1990) ‘Divine Design and the Industrial Revolution: William Paley’s Abortive Reform of Natural Theology’ *Isis* 81: 214-229.

Grafen, A. (1999) ‘Formal Darwinism, the Individual-as-Maximising-Agent Analogy and Bet-Hedging’ *Proceedings of the Royal Society* B 266: 799-803.

Grafen, A. (2003) ‘Fisher the Evolutionary Biologist’ *The Statistician* 52: 319–29.

Grafen, A. (2006) ‘Optimization of Inclusive Fitness’ *Journal of Theoretical Biology* 238: 541-563.

Grafen, A. (2007) ‘The Formal Darwinism Project: A Mid-Term Report’ *Journal of Evolutionary Biology* 20: 1243-1254.

Grafen, A. (2014a) ‘The Formal Darwinism Project in Outline’ *Biology and Philosophy* 29: 155-174.

Grafen, A. (2014b) ‘The Formal Darwinism Project in Outline: Response to Commentaries’ *Biology and Philosophy* 29: 281-292.

Grafen, A. (2015) ‘Biological Fitness and the Fundamental Theorem of Natural Selection’ *American Naturalist* 186: 1-14.

Grafen, A. (2018) ‘The Left Hand Side of the Fundamental Theorem of Natural Selection’ *Journal of Theoretical Biology* 456: 175-189.

Krebs, J. and N. Davies (1997) ‘Introduction to Part Two’ in Krebs and Davies (eds), *Behavioural Ecology: An Evolutionary Approach*, 4th edn., Blackwell Science, Oxford.

Lauder, G. (1996) ‘The Argument from Design’ in M. Rose and G. Lauder (eds.) *Adaptation* San Diego: Academic Press.

Lewens, T. (2004) *Organisms and Artifacts: Design in Nature and Elsewhere*. Cambridge, MA: MIT Press.

Lewens, T. (2005). ‘The Problems of Biological Design’, in A. O’ Hear (ed.), *Philosophy, Biology and Life* Cambridge: Cambridge University Press: 177-191.

Lewens, T. (2007) ‘Adaptation‘, in M. Ruse and D. Hull (eds) *The Cambridge Companion to the Philosophy of Biology*, Cambridge: Cambridge University Press.

Lewens, T. (2009) ‘Seven Types of Adaptationism’ *Biology and Philosophy* 24: 161-182.

Lewens, T. (2010) ‘Natural Selection Then and Now’ *Biological Reviews* 85: 829-835.

Lewens, T. (2018) ‘The Nature of Selection: An Overview’ in R. Joyce (ed.) *The Routledge Handbook of Evolution and Philosophy*. London: Routledge.

Lewontin, R. C. (1978) ‘Adaptation’ *Scientific American* 239: 212-230.

Maynard Smith, J. (1969) ‘The Status of Neo-Darwinism’ *Theoretical Biology*, C. H. Waddington (ed.) Edinburgh: University Press.

Miller, G. (2000) *The Mating Mind*. London: Heinemann.

Okasha, S. (2008) ‘Fisher’s Fundamental Theorem of Natural Selection: A Philosophical Analysis’ *British Journal for the Philosophy of Science* 59: 319-351.

Okasha, S. (2018) *Agents and Goals in Evolution*. Oxford: Oxford University Press.

Okasha, S. and J. Martens (2016) ‘Hamilton’s Rule, Inclusive Fitness Maximisation, and the Goal of Individual Behaviour in Symmetric Two-Player Games’ *Journal of Evolutionary Biology* 29: 473-482.

Okasha, S. and C. Paternotte (2014) ‘Adaptation, Fitness and the Selection-Optimality Links’ *Biology and Philosophy* 29: 225-232.

Ollason, J.G. (1987) ‘Artificial Design in Natural History: Why It’s So Easy to Understand Animal Behavior’ in P. Bateson and P. Klopfer (eds), *Alternatives: Perspectives in Ethology*, Vol. 7, Plenum: New York: 233–257.

Paley, W. (1802) *Natural Theology*. First Edition edited by M. Eddy and D. Knight (2008). Oxford: Oxford University Press.

Pinker, S. and P. Bloom (1990) ‘Natural Language and Natural Selection’, *Behavioral and Brain Sciences* 13: 707-784.

Plutynski, A. (2006) ‘What was Fisher’s Fundamental Theorem of Natural Selection and what was it for?’ *Studies in History of Biological and Biomedical Sciences* 37: 59-82.

Price, G. R. (1972) ‘Fisher’s “Fundamental Theorem” Made Clear’ *Annals of Human Genetics* 36: 129–40.

Reiss, J. (2009) *Not By Design: Retiring Darwin’s Watchmaker*. University of California Press.

Scriven, M. (1959) ‘Explanation and Prediction in Evolutionary Theory’ *Science* 130: 477-82.

Sober, E. (1984) *The Nature of Selection*. Chicago: University of Chicago Press.

Sober, E. (2001) ‘The Two Faces of Fitness’ in R. Singh, C. Krimbas, D. Paul and J. Beatty (eds.) *Thinking About Evolution*, Cambridge: Cambridge University Press.