Life, the universe and everything

Andy Gardner

Abstract The Formal Darwinism project probes the connections between the dynamics of natural selection and the design of organisms. Here, I explain why this work should be of interest to philosophers, arguing that it is the natural development in a long-running scholarly enquiry into the meaning of life. I then review some of my own work which has applied the tools of Formal Darwinism to address issues concerning the units of adaptation in social evolution, leading to a deeper understanding of the adaptation of individual organisms. Finally, I sketch some directions Formal Darwinism to explore beyond the biological sciences, with a focus upon cosmology.

Keywords Anthropic bias · Cosmological natural selection · Design · Group selection · Inclusive fitness · Multi-level selection · Optimization · Selfish gene · Superorganism

The meaning of life

The Formal Darwinism (FD) project probes the connection between the dynamics of natural selection and the design of organisms. Specifically, it captures the dynamics of natural selection in the form of George Price's (1970) equation and it captures the design objective of organisms in the form of an optimization program, and it seeks formal correspondences between these two mathematical objects. Its immediate goal is to frame Charles Darwin's argument—that natural selection explains the design of the living world—in formal terms.

A. Gardner (✉)
School of Biology, University of St Andrews, Dyers Brae, St Andrews KY16 9TH, UK
e-mail: andy.gardner@st-andrews.ac.uk

Alan Grafen expresses some uncertainty as to why this work might be of interest to philosophers. So, my first aim for this commentary is to explain why philosophers *should* be interested. Accordingly, here I provide a sketch reminding the reader that Darwin's discovery—and its ongoing formalization—is the natural development in a long-standing scholarly enquiry into the meaning of life.

For many thinkers throughout history, fathoming the meaning of life has been framed in terms of trying to understand the nature of "God". Two means of doing so have suggested themselves. First, "revealed theology" is the attempt to glean information about God from scriptures and miracles. However, this option was thoroughly and convincingly discredited by David Hume. Second, "natural theology" is the attempt to glean information about God from the natural world; in particular, from its apparent design. Hume advocated this option, but cautioned against attributing any qualities to God that could not be directly inferred from the design of the world.

William Paley championed the natural theology approach. He argued that, although all aspects of the cosmos are undoubtedly designed, this is not at all evident from contemplation of the physical world. For example, although the placement of the stars may be according to some grand plan, this is not obvious to us from our vantage point from within the cosmos itself. Instead, Paley pointed to the abundant evidence of design in the biological world as providing the best avenue for investigating the nature of God. By careful consideration of the exquisite "contrivance and relation of parts" of organisms and their organs, Paley built a strong case for the existence of a highly intelligent designer.

However, this natural theology project collapsed with Darwin's discovery of natural selection. By convincingly arguing that the design of organisms is driven not by an intelligent agent but rather by a mindless, mechanical process, Darwin stripped God of its intelligence. Moreover, by barring the last respectable avenue for gleaning empirical information about God, Darwin rendered it an empirically-vacuous irrelevance. From a scientific perspective, Darwin killed God. It is no wonder that he likened the publication of *The Origin of Species* to confessing to a murder.

It is not an exaggeration to suggest that Darwin's discovery is the most far-reaching scientific, philosophical and theological advance of all time. So, the formal exposition of Darwin's argument should be of interest to anyone who has ever pondered the meaning of life.

Who wields design?

Darwin argued that natural selection drives the adaptation of individual organisms. Accordingly, Grafen's exposition of the FD project has focused upon connections between the dynamics of natural selection and the idea of the individual organism as a maximizing agent. But why should the individual be the privileged unit of adaptation? To address this question requires a broadening of the scope of the FD project to encompass other units of biological organization. Here, I review recent

applications of the tools of FD to the adaptation of genes and groups, and I show how these results deepen our understanding of the adaptation of individuals.

Adaptation of genes?

For some researchers, the gene is the preferred unit of biological adaptation. Richard Dawkins (1976) suggests that the gene is the fundamental unit because it is the selfish agent which underpins the altruism of individual organisms, and he feels that selfishness is more fundamental than altruism. More pragmatically, Austin Burt and Robert Trivers (2006) treat the gene as the fundamental unit because it recovers adaptive explanations for various biological phenomena—such as meiotic drive and genomic imprinting—that yield no adaptive rationale at the individual level, selfish or otherwise.

John Welch and I (Gardner and Welch 2011) have taken a FD approach to capture this gene's eye view of adaptation. This required that we decide what the gene actually is. When being most explicit, Dawkins (1976) defines the gene as a distributed agent comprising every copy of a particular allelic type in the population. However, we found this choice to be problematic, because the allele is an encoding of the gene's phenotype, and hence it more closely resembles an agent's strategy than the agent itself. Instead, we found it appropriate to treat the gene as a physical scrap of DNA—i.e., the token rather than the type—so as to more readily decouple the agent from its strategy. This interpretation of the gene is natural for many applications. Indeed, statements like "An offspring's paternally derived genes have different interests from the offspring's maternally derived genes" (Haig 2002, p. 4) must be referring to genes as tokens, because the designations "paternally derived" and "maternally derived" apply to genic tokens and not to allelic types (Gardner 2014).

We found that this token-sense of the gene-as-maximizing agent view does mathematically link with the dynamics of selection (Gardner and Welch 2011). This is unsurprising, because here the gene is mathematically equivalent to a haploid, one-locus individual, and so the usual connections for the individual-as-maximizing-agent analogy are expected to hold. Consequently, in the context of social interactions, we found that genes are designed to maximize their inclusive fitness, rather than their personal reproductive (or replicative) success. Thus, genes—like individuals—may behave altruistically or spitefully, rather than strictly selfishly (cf. Burt and Trivers 2006). Indeed, the vast majority of genes in the somatic tissues of multicellular organisms do devote their lives to altruism, in order to benefit their relatives in the germ line.

However, a key difference between the gene and the individual is the vast difference in their potential for complex design. A gene has, essentially, only a single trait at its disposal. Hence, it has much less potential to exhibit the contrivance and relation of parts that Paley identified as the hallmark of design. If we consider that a gene may contain several codons—i.e., a "supergene"—then the difference becomes a quantitative rather than a qualitative one. For example, a transposon may be validly considered a single gene, and its molecular behaviour does evidence complex design. However, the impression of design is much less forceful, much less persuasive, than it is at the level of the whole organism.

Adaptation of groups?

The other candidate unit of biological adaptation is the social group. The idea of the adapted group or "superorganism" is an old one within evolutionary biology, but it has enjoyed something of a revival in recent years, buoyed by the growth in interest in multi-level selection. Indeed, Elliott Sober and David Wilson (1998) have argued that adaptation at the group level occurs whenever there is selection at the group level and, hence, since group-level selection is ubiquitous, group-level adaptation is ubiquitous also. However, if adaptation is understood to imply the stronger notion of design, then this argument is not convincing. Moreover, our experience of real-world organisms suggests that group-level design is, in fact, rare.

To address the problem of group-level adaptation, Grafen and I have developed a group-as-maximizing-agent analogy, and have shown that it can be linked to the dynamics of multi-level selection, but only in certain situations (Gardner and Grafen 2009). In particular, we found that what is needed is both the presence of selection between groups and also the absence of selection within groups. This fits well with biological experience. For example, the Portuguese Man-o-War is a classic superorganism—a colony of multiple individuals, that is often mistaken as an organism in its own right owing to its complex group-level design—and it achieves individuality owing to its constituent members being clonally related to each other, which prevents any scope for selection within the colony. In contrast, most groups in most species are genetically heterogeneous, giving ample scope for internal selection, and hence internal conflict. Consequently, social groups are typically not units of adaptation.

General principles of adaptation

I have generalized the above group-adaptation analysis to yield two general principles of adaptation, occurring at any level of biological organization (Gardner 2013; see also Sober and Wilson 2011):

> *Williams' principle* – Adaptation of an entity at any level of biological organization requires selection *between* entities at that level of biological organization.

> *Maynard Smith's principle* – Adaptation of an entity at any level of biological organization requires the absence of selection *within* entities at that level of biological organization.

The first principle echoes George Williams' statement: "Only by a theory of between-group selection could we achieve a scientific explanation of group-related adaptations" (Williams 1966). Williams was railing against the lazy group adaptationism of mid-20th century biologists who interpreted all social phenotypes as being for the good of the group or the species, pointing out that, in order for this to be the case, there would need to be selection operating at this level of biological organization. The second principle echoes John Maynard Smith's statement: "The achievement of individuality at the higher level has required that the disruptive

effects of selection at the lower level be suppressed" (Maynard Smith 1988). Maynard Smith was highlighting the tendency for selection within groups to drive conflicts of interest within the group.

A point can be made here concerning the pitfalls of informal, verbal reasoning about units of adaptation, and the necessity for formal analysis. Sober and Wilson (2011) have advocated the use of a version of Williams' principle to decide the adaptive status of biological entities, and argue that this leads individual organisms to cease being adapted entities when all of selection operates at the group level. Their argument is that selection among individuals is required for individual-level adaptation, and there is no selection among individuals in this context. However, they have interpreted selection among individuals to be synonymous with within-group selection. If instead it is interpreted to refer to the genetic change that is driven by fitness differences among all individuals in the population, regardless of their allocation to different groups, then there is selection among individuals in this context (and it is all happening at the between-group level).

Clearly, whether or not individuals are adapted—that is, show evidence of complex design—should not hinge upon a verbal definition of individual selection. Happily, the FD analysis highlights the quantities that matter. In particular, irrespective of whether or not we call it "individual selection", it is the totality of the selection among all individuals in the population, within and between groups, which the formal analysis pinpoints as driving individual-level adaptation (Gardner and Grafen 2009; Gardner 2013). This may be non-zero even when there is no selection occurring within any group. In such cases, the researcher is free to interpret adaptation as occurring at both the individual and the group level simultaneously. For example, one is free to switch between treating a phagocyte, engulfing an invading bacterium, as an altruistic, inclusive-fitness maximizing agent, on the one hand, versus treating the multicellular organism as launching an immune response against bacterial invaders, on the other hand.

The individual as the privileged unit of adaptation

Consideration of the adaptive status of the gene and the group gives a better appreciation for why the individual organism is the privileged unit of adaptation. First, as remarked upon above, only individuals and groups have sufficiently many traits to show the "relation of parts" that Paley highlighted as a key ingredient of design. In contrast, a single gene will usually encode too few traits for complex design to manifest at this level. Second, only genes and individuals typically enjoy sufficient adaptive integrity to show the "contrivance" that Paley highlighted as the other key ingredient of design. Genes, being unitary, rather trivially show a unity of purpose, and individuals—having mechanisms like clonality that more-or-less abolish internal selection—are more-or-less integrated units of purpose. In contrast, most groups will be riven by internal selection and consequently their members are prone to work at cross purposes. Consequently, only the individual organism emerges as robustly manifesting complex design, which explains her privileged status among units of adaptation (Table 1).

Table 1 The individual is the privileged unit of adaptation

	Scope for complexity?	Adaptive integrity?
Gene	No	Yes
Individual	Yes	Yes
Group	Yes	No

Unlike the gene, the individual has many parts and hence scope for complexity. And, unlike the group, the individual enjoys adaptive integrity owing to the relative absence of internal selection

These points highlight that the study of adaptation is, first and foremost, a whole-organism pursuit. Bearing this in mind provides simple resolution of a variety of philosophical issues that have arisen owing to a narrow focus upon individual traits. For example, Jerry Fodor and Massimo Piattelli-Palmarini (2010) suggest that there is no proper basis for saying that any product of natural selection is *for* any purpose. Their argument runs as follows: we might want to say that a frog's darting tongue is for catching flies, but we might alternatively say that it is for catching "ambient black nuisances"; and, since it is empirically impossible to tease these alternatives apart, we must abandon all hope of interpreting the frog's tongue as being *for* anything. However, the FD approach provides an alternative—mathematical, rather than empirical—avenue for exploring the function of the products of natural selection, which engages with the design objective of whole organisms rather than any one of their parts. The results of the FD project allow us to confidently say that all of the frog's adaptations are *for* maximizing her inclusive fitness.

The purpose of the universe

At the core of the FD project is a formal connection between the dynamics of selection, as captured by Price's equation, and the notion of agency or design, as captured by an optimization program. Grafen's exposition of the FD project has focused upon genetical, natural selection and the idea of biological individuals as fitness-maximizing agents. However, Price (1995) intended for his covariance formula to underpin a general theory of selection, stretching beyond evolutionary biology. Consequently, it is useful to explore potential applications of the FD approach to claims about selection and design in non-biological media. As a vivid illustration of its possible scope, here I discuss the application of the FD project to cosmology.

There is growing interest in the idea that the universe appears as if it has been fine-tuned. The standard model of physics rests upon ~ 30 dimensionless parameters that take seemingly arbitrary values that vary over many orders of magnitude. The physicists' model gives no indication as to why the parameters take their particular values, but it is widely agreed that if they took even slightly different values the universe would look strikingly different (for an in-depth review, see Tegmark et al. 2006; for an opposing view, see Stenger 2011). For example, an

increase in the value of the cosmological constant would lead to the universe expanding too rapidly for stars, galaxies and life to arise (Weinberg 1987).

Perhaps unsurprisingly, this apparent design has led to a reinvigoration of natural theology, and the claim that this fine-tuning is evidence of a supernatural, intelligent designer. However, two alternative, naturalistic explanations have been given to account for this apparent fine-tuning. Both appear to be amenable to FD analysis.

First, Lee Smolin (1992) has suggested that a process of "cosmological natural selection" (CNS) has literally fine-tuned the universe, to maximize its production of black holes. His argument is based on the assumption of a multiverse, comprising many universes that vary in their parameterization of the standard model, and with new universes being born inside black holes and inheriting—with slight mutation—the parameter values of their parent universes. Joseph Conlon and I have framed this argument in FD terms, capturing the dynamics of CNS using Price's equation and capturing the idea that the purpose of the universe is to produce black holes using an optimization program (Gardner and Conlon 2013). The usual mathematical connections obtain, and so the logic of the argument appears to hold. But the analysis highlights deep conceptual problems with the hypothesis, such as what physical interpretation one should attach to between-generation change in the CNS model, given that time is not defined at the scale of the multiverse.

Second, some degree of apparent cosmic fine-tuning may be explained by the "weak anthropic principle" (Barrow and Tipler 1986). That is, the universe might conceivably have been rather different, but it cannot help but appear fine-tuned to us, because if it were any other way we would not be here to observe it. Whether this is a convincing explanation is the subject of on-going debate. Some have argued that although it could explain why the universe is fine-tuned, given that it is observed, it does not explain why the universe is fine-tuned and observed rather than not-fine-tuned and not-observed. Accordingly, some advocates of the weak anthropic principle suggest that it too requires the existence of a multiverse (reviewed by Bostrom 2002).

In Box 1, I sketch what a FD analysis of the anthropic bias effect might look like. Here, Price's equation provides a mapping between the set of possible worlds and the set of worlds that support intelligent observers, capturing the anthropic bias in any characteristic of the world as its covariance with the relative capacity for supporting intelligent observers. And an optimization program captures the corresponding idea that the purpose of the world is to support intelligent observers. Analogues of Grafen's four mathematical correspondences between the dynamics of natural selection and the design of organisms obtain, formalizing the idea that anthropic bias leads the world to appear designed to accommodate intelligent observers.

Here, I have introduced anthropic bias in the context of cosmic fine-tuning. But such observer-selection effects permeate science and philosophy (reviewed by Bostrom 2002), and so their wider connection to FD would be useful to explore. More generally, the conjecture that *all* design-like phenomena are explicable with reference to selection-like phenomena suggests itself. Accordingly, a general proof of the isomorphism between design and selection, in any medium, may be a useful ultimate goal for the FD project.

Box 1 Anthropic bias

I assign every possible world a unique index $i \in I$ and a basic probability π_i. I denote the state of each world by $\tau_i \in T$, and assign each world a real value χ_i according to some arbitrary character of interest. Next, I assume that the number of intelligent observers dwelling in each world is a function of its state; that is, $\omega_i = \Omega(\tau_i)$. Consequently, choosing an intelligent observer at random, each world's probability of containing that observer is $\pi_i' = (\omega_i/\omega)\pi_i$, where I have followed Price's approach of dropping subscripts to denote an expectation: i.e. $\omega = \Sigma_I \pi_i \omega_i$.

The observer effect may be defined as the difference between the conditional and unconditional expectation of the character of interest: i.e. $\Delta\chi = \Sigma_I \pi'_i \chi_i - \Sigma_I \pi_i \chi_i$. Some algebra shows that that this may be expressed as:

$$\Delta\chi = \mathrm{cov}_I(\omega_i/\omega, \chi_i), \tag{B1.1}$$

i.e. it is a selection-covariance effect of the Price (1970, 1972) form.

The idea that the world is designed for the purpose of supporting intelligent observers may be captured using an optimization program. Specifically, I write:

$$\tau \max_{\tau \in T} \Omega(\tau) \tag{B1.2}$$

which formalizes the idea of the world's state τ being chosen so as to maximize the number of intelligent observers dwelling in that world, $\Omega(\tau)$. This allows a formal assessment of a world's optimality with respect to supporting intelligent observers. An optimal world is one that solves the optimization program; that is, it has a state τ^* such that $\Omega(\tau^*) \geq \Omega(\tau)$ for all $\tau \in T$. It is useful to denote the maximal maximand value as $\omega^* = \Omega(\tau^*)$. And a suboptimal world is one that does not solve the optimization program; that is, it has a state τ° such that $\Omega(\tau^\circ) < \Omega(\tau)$ for at least one $\tau \in T$.

The following correspondences obtain;

I. *If all possible worlds are optimal, then there is no scope for anthropic bias.* If the optimization program (B1.2) is solved, then $\omega_i = \omega^*$ for all $i \in I$, and hence, from equation (B1.1), $\Delta\chi = 0$.

II. *If all possible worlds are suboptimal, but equally so, then there is no scope for anthropic bias.* At least one possible world must be optimal, by definition. Thus, the antecedent is false, and this correspondence describes a vacuous truth.

III. *If the possible worlds vary in their optimality, then there is scope for anthropic bias, and anthropic bias in each character is equal to its covariance with relative attained maximand.* From equation (B1.1), $\Delta\chi = \mathrm{cov}_I(\omega_i/\omega, \chi_i)$.

IV. *If there is no scope for anthropic bias, then all possible worlds must be optimal.* Combining correspondences I, II and III yields a biconditional form of correspondence I, which implies the truth of the reverse conditional, i.e. correspondence IV.

References

Barrow JD, Tipler FJ (1986) The anthropic cosmological principle. Oxford University Press, Oxford

Bostrom N (2002) Anthropic bias: observation selection effects in science and philosophy. Routledge, New York

Burt A, Trivers R (2006) Genes in conflict. Harvard University Press, Cambridge

Dawkins R (1976) The selfish gene. Oxford University Press, Oxford

Fodor J, Piattelli-Palmarini M (2010) What Darwin got wrong. Profile Books Ltd, London

Gardner A (2013) Adaptation of individuals and groups. In: Bouchard F, Huneman P (eds) From groups to individuals. MIT Press, Cambridge

Gardner A, Conlon JP (2013) Cosmological natural selection and the purpose of the universe. Complexity 18:48–56

Gardner A (2014) Genomic imprinting and the units of adaptation. Heredity (in press)

Gardner A, Grafen A (2009) Capturing the superorganism: a formal theory of group adaptation. J Evol Biol 22:659–671

Gardner A, Welch JJ (2011) A formal theory of the selfish gene. J Evol Biol 24:1801–1813

Haig D (2002) Genomic imprinting and kinship. Rutgers University Press, New Brunswick

Maynard Smith J (1988) Evolutionary progress and the levels of selection. In: Nitecki MH (ed) Evolutionary progress. University of Chicago Press, Chicago

Price GR (1970) Selection and covariance. Nature 227:520–521

Price GR (1972) Extension of covariance selection mathematics. Ann Hum Genet 35:485–490

Price GR (1995) The nature of selection. J Theor Biol 175:389–396

Smolin L (1992) Did the universe evolve? Class Quantum Gravity 9:173–191

Sober E, Wilson DS (1998) Unto others: the evolution and psychology of unselfish behavior. Harvard University Press, Cambridge

Sober E, Wilson DS (2011) Adaptation and natural selection revisited. J Evol Biol 24:462–468

Stenger VJ (2011) The fallacy of fine-tuning: why the universe is not designed for us. Prometheus Books, New York

Tegmark M, Aguirre A, Rees MJ, Wilczek F (2006) Dimensionless constants, cosmology and other dark matters. Physical Review D 73:023505

Weinberg S (1987) Anthropic bound on the cosmological constant. Phys Rev Lett 59:2607–2610

Williams GC (1966) Adaptation and natural selection. Princeton University Press, New Jersey