



THE PHILOSOPHY OF SOCIAL EVOLUTION

JONATHAN BIRCH

Reviewed by Carl Brusse & Kim Sterelny

The Philosophy of Social Evolution

Jonathan Birch

Oxford: Oxford University Press, 2017, £30.00

ISBN 9780198733058

1 Introduction

It is obvious that some scientific ideas are more powerful than others, for some spawn much productive research. This is still more impressive when the ideas are themselves simple and easily stated. By this criterion, Hamilton's rule of inclusive fitness (Hamilton [1964a], [1964b]), usually formulated with the inequality $rb > c$, rivals $E = mc^2$. Loosely, Hamilton's rule is an inequality that describes the conditions (costs, benefits, relatedness) under which costly 'social' behaviours, such as altruism and spite, can be favoured by natural selection. Such deceptively simple ideas can become enmeshed in confusions, misunderstandings, and rival interpretations, however. There is much scope for interpretation (both empirical and philosophical) with regards to how Hamilton's rule is applied and the relationships between the associated notions of inclusive fitness, and kin and group selection.

Jonathan Birch's *The Philosophy of Social Evolution* offers a comprehensive, level-headed, clear-eyed, and efficient approach towards defining and defending a 'best practice' interpretation of Hamilton's ideas. It articulates and defends a vision of the explanatory scope of Hamilton's programme, while acknowledging and fending off objections and alternatives. At the start of the book, Birch himself states that it is not a textbook nor an introductory text, but rather one long argument for the relevance

and utility of Hamilton's rule in understanding the evolution of cooperation and other social behaviours. This is a sober, pithy, and accurate summary. In fact, the book is so clearly written (including extremely useful, succinct chapter summaries) that it threatens to make the reviewers' job redundant. At the outset of this review, then, let us say that this is philosophy at its best: a clear, informed, and ambitious synthesis of the conceptual, the formal, and the empirical. *The Philosophy of Social Evolution* should be read by anyone interested in the problem of biological cooperation.

2 Overview

Birch structures the book into two parts, dubbed 'foundations' (Chapters 1–5) and 'extensions' (6, 7, and 8). For the purposes of this review, it is easier to divide it into three functionally distinct components instead. Chapters 1, 2, and 3 (and the scene-setting Chapter 0) lay out, contextualize, and defend (formally and in terms of explanatory utility) a specific formulation and interpretation of Hamilton's rule. Chapters 4 and 5 articulate and update several of Birch's important original contributions to this literature: a unified, 'continuum' notion of kin selection and group selection, and a detailed comparative discussion of two formalizations and interpretations of inclusive fitness. Chapters 6–8 then take the theoretical machinery so constructed, and demonstrate its flexibility and utility (via reasonable modification) for application to horizontal gene transfer, cultural inheritance, and characterizing biological individuality.

Throughout, the philosophical arguments are grounded in both the empirical literature (with illustrative biological examples) and in formal, mathematical rigour. Indeed, several of the chapters (for example, Chapters 2 and 5) have sections devoted to mathematically crunchy derivations, or have arguments that turn on highly technical details of the relevant models. These might be daunting for some readers (including one of us), but the formalism is not formalism for formalism's sake. It is not academic virtue signalling. Hamilton's rule is irreducibly mathematical and the mathematical work here (even if much of it can be taken as read) is generally integral to the philosophical issues in focus. The formal expressions of the variables and relationships are simply their most accurate and efficient description. Luckily, while the proofs and derivations are informative, following them through is (for the most part) optional, and a guide to what can safely be skimmed over is included in the introduction of the book.

2.1 Hamilton's rule and the job it can do

In the first chapter, Birch sets up a detailed schema of social behaviour and its moving parts. The social actions of a biological agent are to be evaluated via Hamilton's familiar four-part framework, which is based on how these actions impact the fitness of agents and social partners: mutually beneficial, altruistic (benefiting partners at cost to agents), selfish (benefiting self at cost to partner), and spiteful (harming partner at cost to oneself). Altruism and spite are the curiosities here, from an evolutionary perspective. However, actions proper are not themselves the direct targets of selection. Following Calcott ([2010]), Birch instead attributes fitness-impacts to tasks: the combined effects of actions (perhaps of many agents) rather than actions themselves. Within a given ecological context, tasks confer fitness costs or benefits on agents, and natural selection then acts on the inherited strategies that generate (again, in concert with environmental cues) the social actions that compositionally feed into the tasks, thus completing a causal circle that makes explicit the collaborative context for inclusive fitness. If recent selection has favoured a social strategy, action, or task despite it being at a net direct cost to the agents in question, then we have a genuine case of altruism or spite to explain. This stipulation is needed to avoid, for example, an out-of-equilibrium exploitation being treated as altruism.

When Birch's preferred 'HRG' formulation of Hamilton's rule (first proposed by Queller [1992]) is introduced in Chapter 2,^[1] it arrives by way of a derivation from the Price equation. This is not done merely for its own sake, as the simplifying assumptions required for the derivation to go through have interpretatively limiting implications for the resulting rule. One example here is the idealization that the 'genic environment' is fixed, that is, that there is no meiotic drive or drift, and that dominance or epistatic effects are negligible. More importantly, there is the strict interpretation of variables r , b , and c purely in terms of population statistics—they are properties of populations, not properties of specific individuals, relationships, or one-off

interactions. Cost and benefit in this regard are the differential fitness effects of having the gene or trait in question, and all three variables can be positive or negative.^[2]

The rest of Chapters 2 and 3 is a studious defence of a specific role for this HRG formulation of the rule: a unifying 'organizing framework' for explaining the selection (or otherwise) of social behaviours within populations. The mean breeding value for a social character increases if and only if $rb > c$. By taking positive and negative values for these terms (that is, rb and c), we can carve a graded space of possible explanations: those dominated by direct selection (direct returns from task completion, reciprocity/reward/punishment, pleiotropy/linkage), those dominated by indirect selection (kin discrimination, limited dispersal, and 'outlaw-like' greenbeards and lateral transfer); hybrid explanations (where both direct and indirect effects are possible), and, finally, non-selective explanations. By non-selective explanations, we mean cases where the evolution of costly social behaviours is explained by meiotic drive or some other non-Hamiltonian mechanism. Characterize a population in terms of these four values and the HRG framework will tell you what sort of evolutionary processes might be in play, and whether it might include altruism or spite. This is the basic utility that Birch identifies for the rule and that he uses to defend it against accusations of tautology, predictive limitations, and rival frameworks.

The value of these chapters is in their clear picture of the work Hamilton's rule is supposed to do. Birch sees it as an invaluable conceptual tool rather than a predictive model. It is a framework for triaging and sometimes justifying specific explanatory ecological models rather than a substitute for them. This is entirely appropriate for such a simple, abstract mathematical structure, and the exercise in expectation management is sensible and convincing.

2.2 A deeper dive

Chapters 4 and 5 are somewhat more stand-alone and philosophically bold. In Chapter 4 we get an intriguing and nuanced interpretation of the well-known equivalence results with respect to kin selection and group selection (see, for example, Hamilton [1975]; Grafen [1984]; Kerr and Godfrey-Smith [2002]; Okasha [2016]), and another useful framework for reconceptualizing these two ideas. Chapter 5 makes a similar comparison, but with regard to two ideas that come directly from Hamilton: inclusive fitness and neighbour-modulated fitness.

In discussing group selection, it is standard to distinguish between two forms of group selection. In MLS2, fitness is the fitness of groups: their potential to resist extinction and form new groups. In MLS1, fitness is the productivity of the group: it is a function of the fecundity of the individuals in the group (Damuth and Heisler [1988]; Okasha [2005], [2006]). While no one doubts that MLS2 is theoretically and conceptually distinct from other approaches to the evolution of social behaviour, there is no such consensus about MLS1. Group selection (in the sense of MLS1, where group-level effects of behaviours distribute to influence individual-level fitness) and kin selection (preferential partnering with kin who are likely to share behavioural traits) are both pathways through which apparently individually costly cooperative behaviour can evolve, through the preferential association of co-operators with other co-operators. Intuitively, group and kin selection seem like two different mechanisms; however, the HRG formulation of Hamilton's condition for kin selection to occur is equivalent to that for group selection (derived from the multi-level version of the Price equation). Birch aims to resolve this apparent tension between the intuitive difference between kin and group selection and their apparent equivalence on this analysis. He first distinguishes between formal equivalence and causal equivalence: two different causal processes can be described by the same mathematical generalization, given a suitable degree of abstraction or idealization. But he also identifies several formal nuances: the two models are not quite formally equivalent and MLS1 formal models rely on the simplifying assumption that populations are group structured.

Obviously, group structure (that is, internal integration, external isolation, and stability over generational time) is something that admits of degrees, and populations with a high degree of natural group structure should be more apt for description via group selection than populations where structure is arbitrarily assigned for modelling purposes. Cleaving firmly to a populational approach, Birch represents degree of group structure with graded parameter G , in contrast with a second

parameter, K , which is the degree of genetic correlation between social partners due to kin-related factors. These factors include limited dispersal (leading to an 'accidental' bias toward close-kin interactions) and kin recognition, through which close kin preferentially interact. This deceptively simple approach generates a two-dimensional parameter space in which populations and/or their explanatory models can be situated. Paradigmatic, idealized cases of group or kin selection will have either G or K set high and the other set low (with more realistic cases distributed much more evenly). A well-mixed population where neither occurs will have both G and K low, whereas clonal colonies will have both high. In these 'symmetric' cases, the explanations are genuinely equivalent, and there is no need to choose between them. Birch also considers several intermediate, asymmetric cases, located elsewhere in K - G space. Using this framework, kin and group selection are both formally equivalent and causally distinct, and blend into one another within a space defined by their real-world population features. Again, this is intended as a conceptual framework, with Birch's K - G space following the spatial framework tradition of Mitchell ([2000]) and Godfrey-Smith ([2009]).

The comparison in Chapter 5 between neighbour-modulated fitness and inclusive fitness is more technical, esoteric, and subtle. Hamilton ([1964a], [1964b], [1970]) introduced both concepts as two formally equivalent understandings of social fitness, though his focus was skewed toward inclusive fitness (which is subsequently the better known). Birch's intent here is to defend Hamilton's approach so described, via a better conceptual contrast and reconstruction of the formal argument, and a defence of inclusive fitness as an analytically superior metric or criterion of improvement. This is one chapter that resists succinct summary, and whether it engages the reader will depend on prior interest in the distinction between inclusive fitness and its less popular conceptual sibling. As it largely defends the conceptual status quo (albeit with intriguing sophisticated argument), it is probably best seen as a rounding out of the foundational work rather than a pivotal chapter in the context of the book.

2.3 Applications and extensions

The ambition in the remaining chapters is to extend Hamilton's ideas, and to demonstrate their broader utility in biological theory. Chapter 6 begins with the observation that lateral gene transfer within microbial populations challenges the assumption that genealogical kin relationships can be clearly delineated. As Birch notes, this is not strictly a problem for Hamilton's rule. Relatedness r (in the statistical interpretation) quantifies relative whole-genotype association between agent and donor, and the effective populational average of r (determined by the likelihood that any given agent's interaction partner is non-random with respect to genetic relatedness or similarity) can be influenced by a number of factors. But another assumption was that the value of r between two determinate agents is itself a fixed quantity, which is clearly not the case when an interaction can be a gene transfer that increases genetic similarity. Birch accommodates this with a 'diachronic' modification to the maths of Hamilton's rule, with the intriguing implication that this could favour the selection of prosocial genes, even from a baseline of zero positive assortment. This extension of Hamilton's rule may help to explain the curious commonality of social behaviour in the microbial world (for example, in the form of public goods production in biofilms), and it might even be relevant to one of the early major transitions: the evolution of multi-gene cells. Minimally, it shows the relevance of formal philosophical analysis to science.

Chapter 8 extends Hamilton by building a parallel model of cultural relatedness. We postpone discussion of this chapter to the next section, where we treat it in some detail. Sandwiched in between these two chapters (in Chapter 7) is a more traditional discussion of social evolution and multicellularity, through the lens of Virchow/Haeckel's original notion of the 'cell state'. This chapter engages with familiar themes in this remarkably active research area, in particular with the social perspective on multicellular organisms and their evolution, and the major transitions in evolution literature. Clonal colonies (with very high K) are well suited to a Hamiltonian analysis of cell-to-cell cooperation, and Birch reviews the various lines of argument that attempt to model a transition from clonal colonies to at least proto-multicellularity. However, he also incorporates economic factors when considering trade-offs between functional specialization, redundancy, and population size, drawing parallels between envisioned proto-multicellular complexes and eusocial insect colonies. As Birch pointed out elsewhere, there is a theoretically challenging trade-off between redundancy and benefit. If the collective functionality of a cell cluster or colony depends on a single individual (one ant, say) or a single cell, that functionality is very fragile. If functionality depends on a

subpopulation, no individual in that subpopulation is essential, and hence b , the benefit of its specific contribution, seems negligible. The analysis in this chapter is a development of Bourke's ([2011]) notion of a positive feedback loop between the size of a social group and specialization, and his own work (Birch [2012]). Though speculative, the ambition is to explain the ratchetting up of multicellularity via a process that is highly (and plausibly) ecologically contingent. There are mechanisms that might come in to play to drive evolution toward organismality, but they require conditions (and the costs and benefits that they entail) to be exactly right. We think that Birch's discussion here (which we will also revisit later) is the sleeper of the book in that the significance of these insights has not yet been fully grasped and is potentially very important. Even if the specific feedback loop turns out not to be important, this analysis shows the power of the Hamiltonian framework.

At first, the ordering of these three chapters might seem curious. But while the chapters on microbial and cultural social evolution are very much of a piece at a formal level, thematically there is a clear progression: the more general discussion of the cell state allows a neat pivot to re-apply the formal models for lateral gene transfer to human social organizations. First cell (proto) state, then human (proto) state. As a result, this final section rounds the book off by opening up profound potential applications, while highlighting the common principles linking them.

3 Discussion

In a book of this scope there will always be gaps. For example, at different points in the book there is the appearance of tension between emphases on mere genetic similarity, independent of genealogical connection, perhaps including greenbeards,^[3] and 'kinship dependant sources of relatedness', defined with respect to genealogical relationships. In Chapter 4, kin selection is defined in genealogical terms via the parameter K , with greenbeards explicitly excluded. That is to say, K summarizes the causal contribution to social behaviour-advantage of genealogical correlation (be that behaviourally directed or a contingent by-product of limited dispersal). This is reasonable since K - G space is supposed to discriminate the causal differences between kin selection and (MLS1) group selection, as commonly understood. However, in Chapter 6 the r -boosting lateral transfer of genetic information is folded into the Hamilton apparatus. Birch comments on this (Section 6.5), but treats it as a semantic issue ('is it still kin selection?'). He makes a reasonable hedge here: greenbeard and lateral gene transfer mediated social behaviours are 'at most, marginal cases of kin selection'. For as genetic similarity between individuals at one locus will not be matched at others, we expect intra-genomic conflict to be a significant factor. But given the stipulation that kin selection is to be characterized by shared genealogy, if social behaviour evolves through those genetic similarities, would that be even a marginal case of kin selection? More significantly, given that the K - G space is defined by the statistical interpretation of Hamilton's rule, the choice of a strictly genealogical interpretation of K (and therefore of kin selection) leaves the K - G space looking less impressively comprehensive than it originally seemed, as it now excludes social microbes as well as greenbeards. How should the modelling of lateral gene transfer (which is so convincingly portrayed as ubiquitous) be compared with group selection? The moves here are all perfectly reasonable by themselves, but in aggregate the music they make isn't entirely harmonious. Perhaps the book might have included a more categorical tabulation of these various phenomena and their relationships, once they have been carved up and reconstructed in line with Hamiltonian principles.

However, these are minor quibbles. Two more substantive issues are the extension of the Hamiltonian analysis into cultural evolution and the general positioning of this book in relation to previous, similar work. We now turn to these.

3.1 Cultural relatedness in human evolution

Throughout most of the book, Birch develops his Hamiltonian analysis of evolution targeted on paradigmatically biological phenomena. In the final substantive chapter, he develops a cultural evolutionary analogue of Hamiltonian evolution, exploring the possibility that cultural relatedness can shape evolutionary trajectories in ways structurally parallel to the role of genetic relatedness. In this exploration, he first nominates a potential explanatory target and couples that target to a methodological maxim. Second, he specifies his favoured conception of a cultural variant. Third, he makes the empirical bet that there are sufficiently many identifiable lineages of cultural transmission of cultural variants for this to be important in human evolution, and hence, with respect to specific cultural traits, there are parent-offspring relationships as well as other forms of cultural

relatedness. The idea is that if Carl and Kim have learned to tie their shoelaces from the same teacher, with respect to that trait they are cultural siblings. Finally, he develops and distinguishes two conceptions of cultural evolution. On the basis of these framework assumptions (and a couple of idealizations), he then builds a formal model of cultural relatedness in parallel to those of genetic relatedness, and derives a cultural parallel to Hamilton's rule. We will flesh out this skeleton, and then discuss some of these framing ideas.

3.1.1 The explanatory target

There is a consensus that for the most part in economic games, humans typically do not behave as if they aimed to maximize their individual economic rewards (in ultimatum, dictator, and public goods games). Instead, they seem to have 'social preferences': non-instrumental desires to advance the interests of others, even at some cost to their own economic interests. As Birch notes, variants of the ultimatum game (in particular) have been run in a range of small world, traditional, non-market cultures, and while there is a good deal of cultural variation, the qualitative result is the same. Few agents are economic maximizers. Some scepticism about this conclusion is reasonable. The interpretation that the play in these games shows social preferences depends on the idea that each player knows that the interactions are one-shot and that their own decisions are anonymous, masked from others. That in turn presupposes that in these small world settings, the participants trust the assurances of the mostly white, mostly foreign experimenters and their local collaborators. Do they? Really? Notwithstanding this suspicion, Birch both accepts the consensus that agents typically have social preferences and (more importantly) that these social preferences in part explain the widespread cooperation found in pre-state, pre-institutionalized human social life. Birch thinks these social preferences are a potential explanatory target of a cultural version of Hamiltonian evolution. For one thing, it is very plausible that cooperation in these traditional societies is guided by culturally learned norms and customs, and thus that much of this cooperation takes place amongst cultural kin: individuals whose actions towards one another are shaped by common cultural sources. For a second, Birch thinks that it is a live possibility that having and acting on social preferences has, for each agent, a fitness cost. If that is so, their wide and stable presence poses an evolutionary puzzle. As Birch is careful to make clear, this is only a possibility. Recall that selection acts on strategies not individual acts. Many forms of cooperation are beneficial, so generosity might well be an occasional expression of a strategy that on average returns fitness benefits. So it may well be the case that social preferences have a direct evolutionary explanation, as an aspect of a typically advantageous strategy. But Birch regards the possibility that they are costly as live, and hence it is worth exploring models in which social preferences are explained by cultural rather than genetic evolution. We see that as a fair assessment, though would probably offer lower odds than Birch.

3.1.2 Cultural variants

Birch follows the lead of Sperber, Boyd, and Richerson in developing an atomistic, informational, and cognitive view of cultural variants (see, for example, Sperber [1996]; Richerson and Boyd [2005]). For them, for the most part, culture is information (or mis-information) in peoples' heads. Variants are mental representations—paradigmatically, beliefs and preferences. We think that is the wrong theoretical choice, though fortunately, we think the rest of the theoretical framework is independent of that choice. We think it is better to focus on the products of agents' minds rather than their contents. Cultural variants are best seen as artefacts and types of action. First, if there is cultural transmission, it is the products not the representations that are perceived and copied. That would be a distinction without a difference except that, often, the relationship between representation and product is many-to-one. Very likely, a group of culturally connected agents will have quite different mental representations of a ritual, narrative, or tool, even though the public phenomena are fairly standardized. There need not be a common informational package that they all have internalized. For one thing, what Thag needs to represent about (say) a specific ritual will depend on what else Thag knows: if the ritual shares common elements or common structures with other rituals Thag knows, he needs to represent less about that specific ritual, and the same is true of Thag's relation with, say, an element in his toolkit. The less well-informed Thrug needs a richer representation of the ritual in question. For another, agents do not act in a physical or social vacuum; the environment can cue and scaffold, say, a ritual, so Thag succeeds in producing the standard variant in part because he is prompted by others' actions and the physical scaffolds. He keeps half an eye on Thrug who keeps half an eye on him: there is an unintentional division of informational labour. The content of Thag's mind need not contain an informationally autonomous specification of each element of his production repertoire and Thag's partial

recipe need not be the same as that of his cultural kin, even if their products are appropriately similar. Second, taking cultural variants to be mental representations buys into the extraordinarily difficult issue of determining when two beliefs—say, the belief that drought is caused by sorcery—count as the same belief. For belief (and desire) re-identification is embroiled with the intractable problems of holism. Under what circumstances do Thag and Thrug count as cultural siblings, with respect to their sorcery beliefs? Suppose Thrug but not Thag thinks that Thrung is a sorcerer. This problem arises for artefacts and practices too, but it seems especially intractable with intentional states. Finally, as Celia Heyes ([2018]) points out, it is surely unwise to tie our models of cultural evolution to the ontology of folk psychology. In the light of all this, we suggest conceptualizing cultural traits as the actions and products of agents: cultural extended phenotypes.

3.1.3 Lineages

Straightforward cultural parallels to Hamiltonian evolution, where cultural relatedness effects evolutionary trajectories, would presuppose clear descent lines. There must be lineages of cultural causation, reasonably clearly defined model–novice relationships. Birch argues that it is plausible that there are cultural lineages of this kind. He need not, and does not, claim that every culturally learned trait of some focal individual has identifiable parents. This is wise, because cultural parentage is generally far more complicated. One of us (Sterelny) was once a quite serious chess player. But Sterelny had no identifiable chess parent(s). He learned the moves from his paternal lineage, but the ability to play was moulded by thousands of interactions with hundreds of others. Cultural transmission is often diffuse, with a given individual's cultural trait the result of a multitude of interactions with many agents. In cases like this, there are no identifiable lineages at the level of individual agents. If transmission is diffuse, cultural selection can only act on a population of collectives, and only then if different versions of a cultural trait diffuse within each collective, with differential fitness consequences. But, plausibly, quite a lot of cultural transmission is not diffuse. Sterelny's curry cooking repertoire does have an identifiable cultural parent (to wit, his first girlfriend). The extent to which cultural transmission segregates into identifiable lineages now and in deep time is an open empirical question. That said, we agree with Birch that it is a fair bet that identifiable lineages with fitness consequences are common enough to be important—very likely more so in small scale communities and hence in deep time.

3.1.4 Modes of cultural evolution

Birch identifies two modes of cultural selection. The first, and his main focus, CS1, is natural selection on culturally transmitted traits. In this mode, we understand fitness in the normal way as the reproductive success of biological individuals, but where that success can depend on culturally rather than genetically transmitted traits, and with the further proviso that patterns of cultural inheritance need not be congruent with genetic inheritance. Given this conception, the Hamiltonian idea is that a trait can increase in frequency if Thag, with this trait, acts towards his cultural siblings in ways that reduces Thag's biological fitness, but sufficiently increases the fitness of his cultural kin. The second mode, CS2, presupposes a transformation in the notion of fitness. The cultural fitness of a Thag-like model with respect to a trait depends on the number of agents who have acquired that trait from that model, discounted by the number of other cultural parents those agents have for that trait. CS2 presupposes well-defined cultural lineages, but it does not presuppose that there is only a single parent for each cultural trait.

In terms of this distinction, Birch makes a methodological proposal and suggests a substantive hypothesis. The proposal is that we should develop CS1-type explanations of cultural phenomena only if direct natural selective explanations seem implausible, and we should develop CS2-type explanations only if neither direct nor CS1 explanations seem plausible. The substantive hypothesis, foreshadowed above, is that social preferences have CS1 explanations.

These are the framework assumptions. On the basis of those assumptions, Birch develops a precise formal parallel of cultural kin selection, vindicating the Hamiltonian intuition about CS1. Prosocial but costly behaviour will be positively selected through CS1-type processes only if there is a source of positive cultural relatedness, namely, if cultural kin preferentially interact. If that condition is met, the positive effect of individuals with that prosocial cultural trait on others with that same trait can more than compensate for their direct fitness sacrifice. That said, the formal derivation involves a number of idealizations and one of these is troubling. It assumes, for each cultural trait, all descendants have the same number of ancestors. This is manifestly

not true of many cultural traits. If, for example, the craft of making a hand axe is learned within the family from all the competent artisans in that family, then as the family grows, younger artisans will have more cultural parents than their older siblings, and artisans in different families will have different numbers of parents. The problem is not that the derivation depends on idealizations. Rather, theory suggests that variation in the number of available models (that is, cultural parents) is causally important, making a major difference to transmission (Henrich [2004]; Powell *et al.* [2009]).

As one would expect from the rest of the book, this is a very measured and clear-sighted analysis of one (and to some degree two) forms of cultural evolution driven by selection. Birch has added another cleanly specified mechanism to the cultural evolution toolbox. There is no suggestion that Hamiltonian cultural evolution is the only form of cultural evolution. To the contrary, Birch clearly distinguishes the empirical presuppositions of Hamiltonian cultural selection from the cultural group selection models developed by Boyd, Richerson, and their colleagues (Boyd and Richerson [1985]; Richerson *et al.* [2003]; Boyd [2016]). Understood this way, and with the modest reservations we have noted en route, Birch has developed a conceptual and theoretical vindication of a cultural version of inclusive fitness and shown that its empirical presuppositions are at least plausible in important cases. This is an impressive concluding chapter of an impressive book.

3.2 Position in the literature

When considering how this book builds on what has come before, a useful comparison is with Peter Godfrey-Smith's ([2009]) *Darwinian Populations and Natural Selection*. This is a book that Birch extensively references, and there are obvious ways in which his book is a natural successor to it (or at least a proximate cousin).

The clearest sense in which *Social Evolution* takes *Darwinian Populations* as its starting point is with respect to the populational approach. Godfrey-Smith begins his book by formulating the familiar Lewontin conditions for natural selection in populational terms, laying out five more-or-less orthogonal dimensions of selection-aptness that define a continuum of population types. This framework allows populations to be recognized as marginally, minimally, partially, or paradigmatically Darwinian for different reasons and to different degrees, thus avoiding quibbling about thresholds for 'proper' natural selection. Birch self-consciously takes a continuum approach as well, first to the carving up the space of social behaviours (Chapter 2), and then to the K - G space that allows kin selection and MLS1 group selection to seamlessly bleed into one another.

This dedication to population thinking also extends to the statistical interpretation of Hamilton's rule itself. In his review of *Darwinian Populations*, Daniel Dennett ([2011]) objected to populational primacy, asserting that an agential approach had always been central to natural selection, especially in the sense of the (as he puts it) *cui bono* question: who the beneficiaries are of competition. Godfrey-Smith ([2011]) pushes back against this in a number of ways, but by treating the Hamiltonian terms b and c statistically, Birch shows how a populational approach incorporating selection can operate despite such concerns. Populations are composed of agents, but granular, agent-by-agent accounting of costs and benefits is not required. The average benefits and costs of a social behaviour-trait within a population (along with relatedness interaction bias) selectively explains its evolutionary trajectory. Population-level phenomena can explain social behaviours (altruistic and antagonistic alike). This is not gerrymandering but rather a natural consequence of the HRG formulation of Hamilton's rule, and the fact that relatedness makes no sense outside the context of a population. In this sense, the power of Birch's Hamiltonian approach to social evolution is a vindication of the population-based framework.

It is also worth noting how the two books can be seen as complementary with regard to how they extend their core theoretical apparatus (populational natural selection and populational kin/group selection) to major transitions in cooperation and individuality. For example, they both focus on the evolution of multicellularity. Similar to the continuum of Darwinian populations, Godfrey-Smith outlines another graded dimensional framework where biological populations exhibit degrees of reproductive specialization, reproductive bottlenecks, and overall functional integration (for example, internal division of labour). When these three parameters are set to zero, we find asocial populations like *E. coli* and libertarians, at higher values we find slime moulds, eusocial insect colonies, and (at the highest settings) collections of cells composing multicellular

organisms like ourselves. This is an elegant conceptual framework (and far more nuanced than this cartoon version here). But it is largely a static analysis. To be fair, Godfrey-Smith emphasizes the stabilizing effect of reproductive bottlenecks: the descendants of the germ cells and queen bees are so highly related that, via broadly Hamiltonian reasoning, this 'de-Darwinizes' the lower-level population by suppressing the incentive for its members to break free of the integrated cooperative regime. Likewise, in his discussion of how transitions might occur, Godfrey-Smith invokes Hamilton's rule (including Queller's version of it) alongside group selection and correlated interaction to help explain the conditions under which cooperation can be stabilized.

However, Birch supplies a framework that identifies a potential mechanism that might serve as a driver for the evolution of social populations, upward through Godfrey-Smith's continuum. As already stated, Birch has developed a picture here of a positive feedback loop between specialization, population size, and redundancy. In this sketch, there are ten discrete effects interacting in a complex causal network. Each of them is independently plausible and/or is well motivated by what came before in the book. For example, greater functional specialization of types can power more efficient task completion, but also permits unused functionality to be sacrificed, lowering the production cost for those more specialized agents, allowing increased group size and thereby greater robustness through numerical redundancy. The entire hypothesized network is plausible and potentially quite powerful—at least under conditions where actual fitness advantages accrue to each of these changes in causal organization. As Birch makes clear, there are realistic confounds (active redundancy is expensive and so on). As a result, his hypothesis has the promising combination of power and fragility: it can explain strong evolutionary pressure toward transitions in individuality, but only if the conditions are absolutely perfect, that is, only in rare circumstances. That is actually a selling point for this hypothesis. We know that such transitions generated long-lasting lineages perhaps a dozen times over, but given the number of single-cell lineages that have existed over the last 3.7 billion years, that is an indiscernibly small proportion of their evolutionary trajectories. This hypothesis suggests an answer to why multicellularity is both possible and rare. It certainly moves the debate on from analysis of the problem and stability considerations and into the realm of evolutionary dynamics. So Godfrey-Smith has built the stage, but Birch has put on the first play.

In a sense, then, *Philosophy of Social Evolution* is the Hamiltonian extension of *Darwinian Populations*, and the two could profitably be read in sequence. They are by no means two parts of a single project, but they so firmly complement each other in the same natural philosophy tradition that we can easily imagine (for example) a coherent higher-level course being built around them.

4 Conclusion

To conclude, this is a superb book. It is lucid, well written, sensible, and intellectually valuable, and the most critical parts are accessible enough for anyone willing to do a bit of homework now and then. It is available to most academic subscribers on Oxford Scholarship (and it is coming out in paperback later this year). It is an essential text for philosophy of biology and, more generally, anyone working on the evolution of cooperation.

Carl Brusse
Department of Philosophy
and
Charles Perkins Centre
University of Sydney
Sydney, NSW, Australia
Carl.Brusse@sydney.edu.au

and

References

- Birch, J. [2012]: 'Collective Action in the Fraternal Transitions', *Biology and Philosophy*, **27**, pp. 363–80.
- Bourke, A. F. G. [2011]: 'Principles of Social Evolution', Oxford: Oxford University Press.
- Boyd, R. and Richerson, P. J. [1985]: *Culture and the Evolutionary Process*, Chicago, IL: The University of Chicago Press.
- Boyd, R. [2016]: *A Different Kind of Animal: How Culture Made Humans Exceptionally Adaptable and Cooperative*, Princeton, NJ: Princeton University Press.
- Calcott, B. [2010]: '[Transitions in Biological Organisation](#)', Ph.D. Thesis, Australian National University, available at
- Damuth, J. and Heisler, I. L. [1988]: 'Alternative Formulations of Multilevel Selection', *Biology and Philosophy*, **3**, pp. 407–430.
- Dennett, D. C. [2011]: 'Homunculi Rule: Reflections on *Darwinian Populations and Natural Selection* by Peter Godfrey Smith', *Biology and Philosophy*, **26**, pp. 475–88.
- Godfrey-Smith, P. [2009]: 'Darwinian Populations and Natural Selection', Oxford: Oxford University Press.
- Godfrey-Smith, P. [2011]: 'Agents and Acacias: Replies to Dennett, Sterelny, and Queller', *Biology and Philosophy*, **26**, pp. 501–15.
- Grafen, A. [1984]: 'Natural Selection, Kin Selection, and Group Selection', in J. R. Krebs and N. B. Davies (eds), *Behavioural Ecology: An Evolutionary Approach*, Oxford: Blackwell, pp. 62–84.
- Hamilton, W. D. [1964a]: 'The Genetical Evolution of Social Behaviour, I', *Journal of Theoretical Biology*, **7**, pp. 1–16.
- Hamilton, W. D. [1964b]: 'The Genetical Evolution of Social Behaviour, II', *Journal of Theoretical Biology*, **7**, pp. 17–52.
- Hamilton, W. D. [1970]: 'Selfish and Spiteful Behaviour in an Evolutionary Model', *Nature*, **228**, pp. 1218–20.
- Hamilton, W. D. [1975]: 'Innate Social Aptitudes of Man: An Approach from Evolutionary Genetics', in R. Fox (ed.), *Biosocial Anthropology*, pp. 133–53.
- Henrich, J. [2004]: 'Demography and Cultural Evolution: How Adaptive Cultural Processes Can Produce Maladaptive Losses: The Tasmanian Case', *American Antiquity*, **69**, pp. 197–214.
- Heyes, C. [2018]: 'Enquire Within: Cultural Evolution and Cognitive Science', *Philosophical Transactions of the Royal Society B*, **373**, p. 20170051.

Kerr, B. and Godfrey-Smith, P. [2002]: 'Individualist and Multi-level Perspectives on Selection in Structured Populations', *Biology and Philosophy*, **17**, pp. 477–517.

Mitchell, S. D. [2000]: 'Dimensions of Scientific Law', *Philosophy of Science*, **67**, pp. 242–65.

Okasha, S. [2005]: 'Multilevel Selection and the Major Transitions in Evolution', *Philosophy of Science*, **72**, pp. 1013–25.

Okasha, S. [2006]: *Evolution and the Levels of Selection*, Oxford: Clarendon Press.

Okasha, S. [2016]: 'The Relation between Kin and Multilevel Selection: An Approach Using Causal Graphs', *British Journal for the Philosophy of Science*, **67**, pp. 435–70.

Powell, A., Shennan, S. and Thomas, M. [2009]: 'Late Pleistocene Demography and the Appearance of Modern Human Behavior', *Science*, **324**, pp. 298–1301.

Queller, D. C. [1992]: 'A General Model for Kin Selection', *Evolution*, **46**, pp. 376–80.

Richerson, P., Boyd, R. and Henrich, J. [2003]: 'Cultural Evolution of Human Co-operation', in P. Hammerstein (ed.), *Genetic and Cultural Evolution of Cooperation*, Cambridge, MA: MIT Press, pp. 373–404.

Richerson, P. J. and Boyd, R. [2005]: 'Not by Genes Alone: How Culture Transformed Human Evolution', Chicago, IL: University of Chicago Press.

Sperber, D. [1996]: 'Explaining Culture: A Naturalistic Approach', Oxford: Blackwell.

Notes

[1] The HRG formulation is a slightly more mathematically nuanced formulation of Hamilton's idea that the strategy of helping a relative will increase in frequency if the benefit to the relative exceeds the cost of help weighted by the degree of relationship, in that it makes the importance of variation in social traits in the population explicit.

[2] By way of verification that this is not a completely introductory text, Birch omits here the obvious clarificatory point that $r = 0$ represents the average relatedness of random pairings within the population. So the relatedness of Carl to Kim can be negative if (for example) Kim's population is small and highly inbred, and Carl is a recent immigrant (in such a population, spiteful action against newcomers has a Hamiltonian incentive). In general, the book errs on the side of streamlined abstraction rather than belabouring the introduction of fundamental terminology with 'for example' illustrations.

[3] Greenbeards are a famous evolutionary thought experiment about the evolution of cooperation: a phenotype that combines the strategy of cooperating only with other cooperators, together with an observable signal of playing that strategy, namely, having a green beard. The problem, of course, is to stop non-cooperators from dyeing their beards green.

Share this review

