Hamiltonian inclusive fitness:
a fitter fitness concept

James T. Costa

In 1963–1964 W. D. Hamilton introduced the concept of inclusive fitness, the only significant elaboration of Darwinian fitness since the nineteenth century. I discuss the origin of the modern fitness concept, providing context for Hamilton’s discovery of inclusive fitness in relation to the puzzle of altruism. While fitness conceptually originates with Darwin, the term itself stems from Spencer and crystallized quantitatively in the early twentieth century. Hamiltonian inclusive fitness, with Price’s reformulation, provided the solution to Darwin’s ‘special difficulty’—the evolution of caste polymorphism and sterility in social insects. Hamilton further explored the roles of inclusive fitness and reciprocation to tackle Darwin’s other difficulty, the evolution of human altruism. The heuristically powerful inclusive fitness concept ramified over the past 50 years: the number and diversity of ‘offspring ideas’ that it has engendered render it a fitter fitness concept, one that Darwin would have appreciated.

1. Introduction

William Donald Hamilton (1936–2000) is one of the greatest evolutionary biologists since Darwin, as measured by the sheer number and diversity of studies directly and indirectly inspired by the central insights of his 1963 and 1964 papers [1,2] on social evolution alone: those introducing the concept of inclusive fitness. The evolutionary phenomena that this concept illuminated (altruism, social systems, mating systems, sex ratio theory, life-history strategies, kin recognition and discrimination, selfish genetic elements, genomic imprinting, etc.) reflect the diversity of taxa on which it has inspired study: over 700 papers encompassing all three domains and five kingdoms of life were tagged with ‘inclusive fitness’ between 1969 and 2013 (ISI/Web of Knowledge search, April 2013). The semicentennial of Hamilton’s landmark papers provides an opportunity to celebrate the discovery of inclusive fitness. Hamiltonian inclusive fitness and such related concepts as kin selection, altruism, reproductive division of labour and eusociality are so much a part of the lexicon of modern biology that it is easy to overlook the relative recency of its origin. Indeed, it is unappreciated among many students of social evolution that the concept of inclusive fitness introduced in 1963–1964 represents the only significant extension of the fitness concept since its derivation in the mid-nineteenth century—an insight that took nearly a century after that of fitness itself.

Disagreement over the definition of fitness and the mode of action of natural selection persists [3–6]. My purpose is not to review fitness concepts or their history, but to provide a general overview of inclusive fitness within the broader fitness concept framework. Accounts of the development of inclusive fitness are available [7–9], but none considers Hamilton’s innovation in relation to the deeper history of fitness and earlier, more general expressions of the indirect fitness idea. Accordingly, I here first offer a concise overview of the origin and evolution of Darwinian fitness, and then discuss Hamiltonian inclusive fitness as an elaboration of this concept that at once solved Darwin’s ‘special difficulty’—caste polymorphism and sterility in social insects—and shed light
on Darwin’s puzzlement over human altruism. In the process, I discuss how Hamilton’s insights renewed the focus on individuals rather than groups as the units of selection while shedding light on genic-level selection.

2. A fitting term

Contrary to the popular belief that Darwin introduced the term and concept of fitness in his landmark *On the Origin of Species* [10] in 1859, this is not the case. While Darwin certainly saw his mechanism of natural selection ending individuals with traits (adaptations) that promoted their survival and reproduction, in the *Origin* he wrote of organisms becoming, through selection, ever-better fitted to the demands of their environment. He did not use the word ‘fitness’ as a descriptive state, nor the words ‘fit’ or ‘fitness’ in the current biological sense of survival and reproductive success. In fact, ‘fitness’ appears just once in the *Origin*, in the final, concluding, chapter: ‘Nor ought we to marvel if all the contrivances in nature be not . . . absolutely perfect; and if some of them be abhorrent to our ideas of fitness’ [10, p. 472]. Here, Darwin uses fitness in the sense of a state of being well adapted. ‘Fit to environment’ is the context in which all cognates of ‘fitness’ are used in the *Origin* (table 1), usage conforming to the natural theology tradition of the time [5].

Five years after the publication of the *Origin*, Herbert Spencer coined the phrase ‘survival of the fittest’, which he equated with natural selection: ‘the survival of the fittest must nearly always further the production of modifications which produce fitness; whether they be modifications that have arisen incidentally, or modifications that have been caused by direct adaptation’ [11, p. 455]. Thus Spencer’s formulation gave birth to a new way of conceptualizing that which natural selection favours: a new entity called ‘fitness’, which is produced by or results from selection favouring traits that confer superior survival and reproductive success, and which *de facto* are passed on in reproduction. Indeed, the Oxford English Dictionary defines biological fitness, ‘fitness to survive and reproduce’, in terms of Spencer’s framing of Darwin’s concept. Alfred Russel Wallace, co-discoverer of evolution by natural selection, was the first to urge Darwin to adopt Spencer’s phrase ‘survival of the fittest’ in place of natural selection, objecting to the implication of conscious intentionality inherent in the word ‘selection’ [12]. Darwin did coin natural selection with human-mediated (and thus conscious) plant and animal breeding in mind, but while he defended his choice of terms, in the fifth (1869) edition of the *Origin* he added Spencer’s phrase as a synonym for natural selection.

The idea of fitness as a state of organisms that may change or be optimized over time by selection was not mathematically formalized until the early twentieth century with the rise of population genetics, combining the approaches of the biometrician school (quantitatively studying variation) with Mendelian genetics [13,14]. Pioneering evolutionary geneticists Fisher, Haldane and Wright developed early mathematical treatments of fitness. In perhaps the first such treatment, Fisher [15] described the fitness of Mendelian genotypes relative to one another in the context of conditions for equilibrium under selection. Fisher later defined Darwinian fitness in terms of the relative number of offspring produced by individuals of a given genotype, based on his ‘Malthusian parameter’—the rate of change of a population over time [16]. Fisher also showed that the rate of increase in fitness owing to natural selection equals the genetic variance in fitness—his Fundamental Theorem of Natural Selection. The Fundamental Theorem is viewed as overly simplistic today [17,18], but Fisher’s treatment of average population fitness and its variation was an important milestone.

Another milestone for a more biologically realistic expression of fitness stemmed from Haldane [19], in the lead paper of his series *A Mathematical Theory of Natural and Artificial Selection* published between 1924 and 1934. Haldane [20] and Norton [21] then treated selection with overlapping generations (unlike the discrete-generation models of Fisher and others) and showed that the intrinsic rate of population increase (Fisher’s Malthusian parameter) is a fitness measure under such conditions as random mating and weak selection. ‘Relative fitness’, growing out of the work of Wright [22], is a common expression of fitness today. Here, the fitnesses of different genotypes of a locus (measured as lifetime reproductive success) are scaled relative to the one with the greatest absolute fitness. In practice, however, assigning a genotypic fitness standard can be problematic because differences in mean relative fitness are defined within and not compared among populations. ‘Standardized relative fitness’, defined in terms of relative rates of increase and a scaling of fitness to the mean fitness of a locus and not one of its genotypes, offers a solution to this problem [23, pp. 166–170].

3. Towards a fitter fitness concept

A concept of fitness that invoked differences among individuals was accepted through the mid-twentieth century. An alternative concept arose that invoked ‘good of the species’ thinking, where groups, populations or species as a whole were seen as the unit of selection, leading to further refinement of the fitness concept. The group-level selection view was most visibly and effectively expressed by Allee et al. [24] and Wynne-Edwards [25]. Wynne-Edwards’ book in particular had a galvanizing effect on Fisherian evolutionists [26–28], who showed group selection to be an unnecessary or incorrect alternative to individual selection and inspired an explicitly genic-level selection view of the evolutionary dynamic most prominently expressed by Dawkins [29].

Altruism posed a problem for the classical view of selection and fitness. Hamilton later expressed surprise that the inclusive fitness concept had been overlooked by early evolutionary

<table>
<thead>
<tr>
<th>term</th>
<th>no. occurrences</th>
<th>pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>fitness</td>
<td>1</td>
<td>p. 472</td>
</tr>
<tr>
<td>fit</td>
<td>2</td>
<td>pp. 158, 183</td>
</tr>
<tr>
<td>fittest</td>
<td>0</td>
<td>n.a.</td>
</tr>
<tr>
<td>fitting</td>
<td>5</td>
<td>pp. 66, 80, 83, 224, 471</td>
</tr>
</tbody>
</table>
geneticists, some of whom touched on the altruism problem but then skirted the central issue. Fisher [16] briefly addressed altruism in the contexts of human society and the evolution of apomatisms and distastefulness in larval insect colonies, which he assumed were sibling groups: ‘The selective potency of the avoidance of brothers will of course only be half as great as if the individual itself were protected . . . ’ [16, 2nd edn., p. 177]. Haldane [30,31] expressed this idea more fully, recognizing that it also applied to extended families with progressively lower degrees of genetic relationship:

Let us suppose that you carry a rare gene which affects your behaviour so that you jump into a river and save a child, but you have one chance in ten of being drowned, while I do not possess the gene, and stand on the bank and watch the child drown. If the child is your own child or your brother or sister, there is an even chance that the child will also have the gene, so five such genes will be saved in children for one lost in an adult. If you save a grandchild or nephew the advantage is only two and a half to one. If you only save a first cousin, the effect is very slight. If you try to save your first cousin once removed the population is more likely to lose this valuable gene than to gain it. [31, p. 44]

Two years later Williams & Williams [32] also addressed the evolution of altruism, though they explicitly avoided this word, preferring instead ‘social donors’ as an expression with less baggage.

Hamilton acknowledged Fisher and Haldane in his 1963 and 1964 papers, and the first chapter of his 1996 collected papers compendium is entitled ‘Shoulders of Giants’ as homage to them as well as to Wright, whose coefficient of relationship [33,34] provided Hamilton with the key term for genetic relatedness that he sought for his indirect fitness formulation. This intellectual thread has special resonance in that Wright developed his approach studying livestock pedigrees, and thus domestication played a role in both Darwin’s initial insights into the evolutionary process [35,36] and in Hamilton’s elaboration of Darwinian fitness a century later.

Hamilton’s first publication on the evolution of altruism—i.e. the first place in which he discussed the inclusive fitness concept—appeared in 1963, though this short paper was written and published while the manuscript that became the paired 1964 papers wended its way through a slow review and revision process [7, pp. 3; 29]. The word ‘inclusive’ appears twice in the 1963 paper in adjacent paragraphs [2, p. 355], reflecting the concept of inclusive fitness even though it does not use the term. In the first instance, Hamilton discussed altruistic behaviour that adds to the ‘genotype-reproduction’ of the actor (altruistic individual), stating that by this he meant the genotype ‘inclusive of the reproduction of identical genes’ in relatives. In the second instance, he was explicit that ‘fitness is reckoned in terms of ‘inclusive’ genotype-reproduction’. This followed his statement that ‘if the gain to a relative of degree r is k-times the loss to the altruist, the criterion for positive selection of the causative gene is k > 1/r’. This relationship, which became known as Hamilton’s rule, was later expressed more intuitively as Br – C > 0, or Br > C (where B and C are the respective benefits and costs conferred on an actor, weighted by the degree of relatedness (r) between the actor and the recipient of the actor’s assistance).

Hamilton later came to view his initial attempts at expressing inclusive fitness rather clumsy. To his delight, his results were derived via a very different approach by Price, who used the covariance between genetic traits and fitness [37,38] to generalize inclusive fitness theory, an advance that Hamilton found at once elegant and more informative. Price’s approach showed the mathematical equivalence of Hamilton’s kin-based formulation and group selection; indeed, the former proved to be simply a special case of the latter [39–41], differing only in the way in which the genetic variance is partitioned. This fact surprised Hamilton at first, but he soon saw the power of Price’s approach, leading him to rethink the idea of group selection. This found expression in Hamilton’s tentative if provocative answer to Darwin’s question about the origin of human altruism, discussed below.

4. Darwin’s difficulties

Darwin grappled with the puzzle of altruism in not one but two sorts of organisms—social insects and people—though it should be pointed out that in discussing caste and sterility in social insects he did not frame the problem in terms of altruism per se, which is a modern perspective on the issue. Rather, Darwin wondered how caste polymorphism and sterility could arise by natural selection, a mechanism that acts to increase survivorship and reproductive success of individuals. He connected morphological divergence among castes to division of labour within families, thus extending the ‘physiological division of labour’ concept of Milne-Edwards [42] as well as his own idea of the ‘ecological division of labour’ which lay at the heart of his principle of divergence [10, pp. 111–126]. The common denominator is that division of labour increases overall efficiency or productivity.

Caste sterility and altruism imply an extreme form of division of labour, however. How could selection favour traits that reduce personal reproduction? Darwin asks in Natural Selection, the forerunner of the Origin, ‘how it is possible that communities of insects should come to possess sterile females or neuters?’ [43, p. 365]. In his solution, Darwin invoked colony (family)-level selection: ‘... it seems not improbable, owing to the vast fecundity of the lower animals, that a certain number of females ... without any waste of time or vital force from breeding, might be of immense service to the community. If this were so ... then natural selection would favour those communities, in which some of the individuals [had been rendered] in some slight degree less fertile than the other individuals’ [43, p. 366]. In the Origin, Darwin posited that the great difficulty posed by morphological divergence and functional sterility ‘is lessened, or, as I believe, disappears, when it is remembered the selection may be applied to the family, as well as to the individual ... ’ [10, p. 237]. He then drew an analogy with domesticated plants and animals. Some individuals are harvested and consumed, but the variety is nonetheless propagated, and even widely spread through preservation and dissemination of the seed stock, a form of group-level selection. Thus, a well-flavoured vegetable is cooked, and the individual is destroyed; but the horticulturist sows seeds of the same stock, and confidently expects to get nearly the same variety; breeders of cattle wish the flesh and fat to be well marbled together; the animal has been slaughtered, but the breeder goes with confidence to the same family [10, pp. 237–238]. The seed stock of Darwin’s ‘well-flavoured vegetable’ is thus analogous to a family group, where certain individuals are selected to propagate the variety, while relatives are destined to be consumed and will never set seed.

Inclusive fitness is, of course, the solution to Darwin’s difficulty over social insects: altruistic (sterile) individuals are
thought to enjoy a net positive inclusive fitness, with indirect fitness gain outweighing their direct fitness loss. Hamilton saw that inclusive fitness works via genic selection, not family-group selection. He parsed genic fitness into personal or ‘classical’ [direct] and ‘neighbour-modulated’ [indirect] fitness components. The indirect component is most often conceived in the context of close relatives, leading Maynard Smith to coin the term ‘kin selection’ [44], yet Hamilton stressed that the inclusive fitness concept is more general than kin selection [45]. Hamilton’s rule is formally modelled in terms of the genetic association (regression) between socially interacting individuals rather than pedigree relatedness ([41], boxes 2 and 3 in [38]).

The other form of altruism that troubled Darwin, that of humans, was addressed in the Descent of Man [46]. In that work, Darwin puzzled over the apparently uniquely human propensity for self-sacrifice, especially for unrelated individuals. Human behaviour is marked by a perennial tension between pro- and anti-social tendencies: ‘The social instincts, which must have been acquired by man in a very rude state, and probably even by his early ape-like progenitors, still give the impulse to many of his best actions’, Darwin wrote, ‘but his actions are largely determined by the expressed wishes and judgment of his fellow-men, and unfortunately still oftener by his own strong, selfish desires’ [46, p. 86]. It may be easy to see where aggression comes from, but how to understand the human propensity for unhesitating self-sacrifice was Darwin’s dilemma. What do we make of those in whom ‘courage and sympathy’ are well developed; those who ‘disregarding the instinct of self-preservation, instantaneously plunged into a torrent to save a drowning fellow-creature’ [46, p. 87]. Darwin had no definitive solution, but seemed to invoke group-level selection, the ‘good of the community’. He concluded after a long discussion that ‘finally, the social instincts which no doubt were acquired by man, as by the lower animals, for the good of the community, will from the first have given to him some wish to aid his fellows, and some feeling of sympathy…’ [46, p. 103].

In his 1975 paper, Hamilton [45] found not only the first expression of Hamilton’s use of Price’s multi-level covariance approach, but a bold application of this to the evolution of human altruism. Hamilton addressed what he termed ‘tribal facies of social behaviour’, offering a response to Darwin’s dilemma by postulating that human social origins occurred in a constellation of small tribal groups with limited inter-group migration, in which tools and language played a key role in conferring disproportionate benefits to groups as a function of intragroup cooperation. This idea lies at the heart of later models of group selection [47–49], which are very different in character from the mid-twentieth century brand of group selection in both the details of the posited selection dynamic and the mathematical rigour of the models.

5. Conclusion

Hamilton concluded his 1975 speculations on human altruism with a discussion of reciprocation, and that provides a fitting segue to my conclusion. While the inclusive fitness concept shed light on the evolution of the kinds of altruism that troubled Darwin—the insectan sort, in the form of division of labour and sterility, and the human sort, in the form of self-sacrifice for non-relatives—it proved to have a far wider currency, and thus gave rise to or complemented with a host of new evolutionary disciplines and subdisciplines. Reciprocal altruism, for example (which Hamilton preferred to call reciprocation), introduced by Trivers [50], nicely complements Hamiltonian inclusive fitness theory; as Dawkins put it, inclusive fitness ‘is the key to understanding half the altruistic cooperation in nature’, while the key to the other half ‘is reciprocation among unrelated individuals’ [51, p. 6]. Reciprocal altruism and the game theoretic approaches it inspired thus join inclusive fitness theory as pillars underlying the modern understanding of social evolution [50,52–54].

Hamilton’s inclusive fitness concept, now 50 years old, has continued to inspire new vistas in the evolution of social behaviour and evolutionary dynamics, from kin recognition and discrimination [55–57] to game theory [58–60] and selfish genetic elements [61,62], and from parent–offspring conflict [63,64] to the brave new world of genomic imprinting [65,66]. Despite lingering problems with the term ‘fitness’ and the limitations of inclusive fitness in some contexts relative to direct fitness approaches [67–72], Hamiltonian inclusive fitness is a fitter fitness concept as judged by the number and diversity of ‘offspring ideas’ it has engendered. It certainly would have intrigued and delighted Darwin for the abundance of light it has shed on the weird and wonderful adaptations of organisms that he and Hamilton, naturalists both, so deeply appreciated.

Acknowledgements. I thank the editors, Joan Herbers and Neil Tsutsui, for the opportunity to contribute to this Special Feature dedicated to Bill Hamilton. Many thanks to librarians Sonja Grund, Anja Brockmann, Marianne Buck and Kirsten Graupner of the Wissenschaftskolleg zu Berlin for their able assistance, to Leslie Costa, Bruce Kogut, Robert Trivers, Raghavendra Gadagkar and two anonymous reviewers for helpful comments and criticisms on the manuscript, and to Joan Herbers for her skillfully, if mercilessly, wielded editorial pen.

Funding statement. This paper was written while a Fellow of the Wissenschaftskolleg zu Berlin; I gratefully acknowledge the support of the Wissenschaftskolleg and Western Carolina University’s Scholarly Leave programme that made my stay in Berlin possible.

References
