BioScience Advance Access published October 6, 2015

Inclusive Fitness Theory Becomes an End in Itself

Social Evolution and Inclusive Fitness Theory: An Introduction. James A. R. Marshall. Princeton University Press, 2015. 196 pp., illus. $39.95 (ISBN: 9780691161563).

Why do some birds raise the offspring of others rather than starting their own nests? Why do some microbes pay a metabolic cost to help each other scavenge iron? Understanding how such social behaviors emerge from natural selection has been an important problem since Darwin first developed the theory of evolution. Of particular interest is the evolution of costly cooperative behaviors, in which an individual sacrifices its own fitness (lifetime reproductive success) to benefit others.

W. D. Hamilton (1964) proposed to analyze these questions by introducing a quantity called *inclusive fitness*. Inclusive fitness modifies an individual's fitness by first removing any amounts of benefit or harm done by others and then adding amounts of benefit or harm done *to* others, with the latter amounts weighted by the degree of relatedness to the recipient. He showed that under certain simplifying assumptions, evolution selects for behaviors that yield the largest inclusive fitness.

To James A. R. Marshall, a professor of theoretical and computational biology at the University of Sheffield and a champion of the theory, "Inclusive fitness theory is probably the most important advance in our understanding of evolution since 1859" (p. xiii). However, the definition of inclusive fitness implicitly assumes that an individual's fitness can be partitioned into separate amounts due to itself and to each other individual. Many (including myself) have asked whether such a partition makes sense for the complex, nonlinear social processes that are ubiquitous in nature (Karlin and Matessi 1983, Nowak et al. 2010, Allen et al. 2013).

To avoid this problem, most works of inclusive-fitness theory look at simplified models in which individuals contribute well-defined amounts of fitness to each other. But this will not do for Marshall, who wants to argue that inclusive fitness is a universal explanation for the evolution of all social behaviors. Instead, following Queller (1992) and Gardner and colleagues (2011), he attempts to attribute portions of an individual's fitness to itself and others in retrospect, using linear regression. To apply this method, the outcome of selection must already be known. One then performs a multivariate linear regression of each individual's fitness on the basis of its own genotype and the genotypes of its interaction partners. The partial regression coefficients are then interpreted as the fitness amounts due to self and others. Applying this method to various game-theoretic models, Marshall concludes that inclusive-fitness theory can "explain the evolution of arbitrary social scenarios" (p. 132).

This method is akin to drawing a best-fit line through a parabola and concluding that linear equations are a universal approach to studying gravity. Although the procedure does not violate any rule of algebra, I cannot identify any scientific question that it answers. Certainly, it does not predict anything, because the outcome of selection must be known before regression can be applied. It also does not reveal why a trait is selected, because correlation does not imply causation. I have found simple examples for which this method completely mischaracterizes the nature of a trait (Allen et al. 2013). For example, a "nurse" who helps the weakest is mislabeled as a "bully," because interacting with the nurse has a negative correlation with fitness. Moreover, the entire procedure is unnecessary: Any relevant question about Marshall's game-theoretic models can be analyzed directly using the tools of differential equations.

Marshall gives a brief nod to the problem of correlation versus causation but defends his method in part by linking it to the field of quantitative genetics. However, in quantitative genetics, linear regression is used either to make inferences from data (e.g., Lande and Arnold 1983) or as an algebraic step in calculating empirically relevant quantities (e.g., expected trait correlations among relatives; Fisher 1918). Marshall, in contrast, uses regression only to rewrite already-known quantities into a form that purports to show inclusive-fitness theory at work.

Why does Marshall (along with others—e.g., Gardner et al. 2011) champion this regression method? Normally, a method without predictive or explanatory power would not survive long in the scientific literature. In this case, however, the method serves a different purpose: to defend inclusive-fitness theory against the criticisms (e.g., Karlin and Matessi 1983, Nowak et al. 2010) that its assumptions are too restrictive. The regression method is designed to show that inclusive-fitness theory is universal, not dependent on any modeling assumptions. But without modeling assumptions, it is impossible to obtain any meaningful conclusions. Instead, inclusive-fitness theory becomes an end in itself, existing only to assert its own truth. Any potential insights the theory might offer are sacrificed in a misguided quest for complete generality.

The vacuity of Marshall's method is most evident in his analysis of "greenbeard" cooperation. Greenbeard cooperators identify each other through a phenotypic marker (say, a green beard) and cooperate only with each other. Initially conceived as a thought experiment, greenbeard genes have been discovered in several species. Marshall shows how different variants of the regression method yield conflicting explanations for how greenbeard genes evolve. One variant finds that greenbeard genes are costly to their bearers and evolve because of high relatedness; another finds that they are beneficial to their bearers and that relatedness is zero. Marshall does not make clear what biological insight we are supposed to draw from these conflicting explanations. His only takeaway message seems to be that inclusive-fitness theory "works," even if its results are self-contradictory.

Marshall's argument culminates with the astonishing claim that "inclusive fitness comes out as the only causal explanation for observed social evolution" (p. 103). In fact, the opposite is true in that the methods that Marshall presents are uniquely *unsuited* to address questions of causality. Causality is typically assessed in a modeling context by varying an empirically relevant parameter and observing the consequences. Marshall presents no such analysis; instead, he proposes to assess causality using yet more forms of linear regression. In other words (and despite his earlier caveats), Marshall's entire argument rests on equating correlation with causation, in defiance of a full century of statistical practice.

Although I have focused so far on what is in Marshall's book, it is at least as striking for what it leaves out. Population genetics, which underlies all of modern evolutionary theory, is given only the briefest of mentions. There is no discussion of how to model different forms of population structure (space, groups, colonies, etc.). Even the mathematically rigorous formulations of inclusive-fitness theory (e.g., Rousset and Billiard 2000) are mentioned only in passing.

Marshall is at his best when describing social behaviors themselves. Certainly, behaviors such as stalk formation in *Dictyostelium discoideum* present fascinating evolutionary puzzles, and Marshall discusses these in a clear and engaging manner. Ultimately, however, his examples only underscore the need for rigorous, empirically grounded mathematical modeling tools. One would not know it from Marshall's book, but such tools already exist in the fields of population genetics, evolutionary dynamics, and evolutionary game theory. Already, these tools have yielded many discoveries, and we can undoubtedly expect many more in decades to come.

References cited

Allen B, Nowak MA, Wilson EO. 2013. Limitations of inclusive fitness. Proceedings of the National Academy of Sciences 110: 20135–20139.

Fisher RA. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Transactions of the Royal Society of Edinburgh 52: 399–433.

Gardner A, West SA, Wild G. 2011. The genetical theory of kin selection. Journal of Evolutionary Biology 24: 1020–1043.

Hamilton WD. 1964. The genetical evolution of social behavior I & II. Journal of Theoretical Biology 7: 1–52.

Karlin S, Matessi C. 1983. The eleventh RA Fisher memorial lecture: Kin selection and altruism. Proceedings of the Royal Society of London B 219: 327–353.

Lande R, Arnold SJ. 1983. The measurement of selection on correlated characters. Evolution 37: 1210–1226.

Marshall JAR. 2015. Social Evolution and Inclusive Fitness Theory: An Introduction. Princeton University Press.

Nowak MA, Tarnita CE, Wilson EO. 2010. The evolution of eusociality. Nature 466: 1057–1062.

Queller DC. 1992: A general model for kin selection. Evolution 46: 376–380.

Rousset F, Billiard S. 2000. A theoretical basis for measures of kin selection in subdivided populations: Finite populations and localized dispersal. Journal of Evolutionary Biology 13: 814–825.

BENJAMIN ALLEN
Ben Allen is affiliated with the Department of Mathematics at Emmanuel College, in Boston, MA, and with the Program for Evolutionary Dynamics and Center for Mathematical Sciences and Applications, Harvard University, Cambridge MA.

doi:10.1093/biosci/biv129