The agent concept is a scientific tool

Samir Okasha: Agents and goals in evolution. Oxford: Oxford University Press, 2018, xiv+254pp, £30.00 HB

Andy Gardner[1]

Okasha (2018) asks whether living organisms and other biological entities can meaningfully be regarded as agents with goals. To answer this question, he surveys a range of different notions of agency—from goal directedness to behavioural flexibility to adaptedness—and discusses the links between the action of natural selection and rational behaviour. His book is thought-provoking, and it provides an excellent entry point into an interesting multidisciplinary literature. I will certainly make use of it in the future as a reference work.

However, I can't help but feel that Okasha has missed an opportunity to properly explore how evolutionary biologists actually use the concept of agency in their research. This is unfortunate, because in science it is the usefulness of a concept, rather than its philosophical tidiness, that provides its ultimate justification. More specifically, whilst Okasha's treatment of agency focuses on *description*, i.e. exploring the extent to which the concept of agency characterizes living organisms, of more interest would be the agent concept's role in *prediction*, i.e. understanding the extent to which an agential perspective facilitates the formulation of scientific hypotheses that can be put to the empirical test.

Okasha embarks on his discussion of agents and goals in evolution by suggesting that the widespread occurrence of these concepts within the evolutionary biology literature is odd from a wider scientific perspective. He remarks that "it is not generally a useful strategy in science to treat the objects of one's study as if they had certain attributes which in fact they lack" (2). But a moment of reflection shows that this is wide of the mark.

Biologists studying population cycles of, say, rabbits often find it convenient to treat population size as if it were a complex number with both real and imaginary components, with the real component being as much a fiction as the imaginary insofar as it allows for fractional (or, indeed, irrational) numbers of individuals. In

✉ Andy Gardner
andy.gardner@st-andrews.ac.uk

1 School of Biology, University of St Andrews, St. Andrews, UK

characterizing the stability of these populations, biologists also make use of the concept of the infinitestimal, which has no literal real-world counterpart.

More generally, all of scientific understanding rests upon models, and all models are false. Models regularly attribute to the natural world features that it does not possess, such as perfectly straight lines, point masses and infinitely distant light sources. Philosophical analysis of these imaginary features entirely misses the point if it does not consider how these ideas are actually used and whether our scientific understanding is richer on account of us having used them.

To the extent that Okasha does touch upon evolutionary biologists' use of the concept of agency, he paints an unfamiliar picture of the researcher identifying a particular feature of a particular organism and then agonizing over whether or not that feature is an adaptation and how closely it approaches its optimal state. I do not recognize my own work in this caricature, nor do I recognize the work of my colleagues. Evolutionary biologists spend their time pursuing more interesting and scientifically productive programmes of research.

Evolutionary biologists are not generally in the business of identifying adaptations. We are usually more interested in explaining variation, be it between genes, individuals, populations, or species. And we do this by developing formal or informal theoretical models, using the models to derive logical predictions, and testing the predictions against empirical data.

A hugely successful way of conducting this kind of work is through application of the optimality approach (Parker and Maynard Smith 1990), often coupled with the comparative method (Harvey and Pagel 1991). The optimality approach helps evolutionary biologists narrow the space of all possible trait variants down to particular, concrete predictions, by identifying trait optima. This is where the concept of agency comes in. If we are to think of a trait as *having* an optimum, which is different from thinking of the trait *being* at this optimum, then we must have in mind some agent with some goal, such that the optimum is defined as what best realizes the agent's goal.

The concept of agency facilitates the formulation of theoretical models, by which I mean any hypotheses about how the world of evolutionary biology works. Okasha points out that humans are particularly fond of thinking in terms of agency, but what he does not mention is how *good* we are at solving problems when they are framed in this way. Logical problems that stump most people are rendered immediately tractable when reframed in terms of agents with goals. Accordingly, evolutionary biologists are more likely to notice interesting patterns or anticipate important effects when we are free to think about living organisms as agents with goals than if we are forced to restrict our attention to the dynamics of genotype frequencies or some other non-agential mode of analysis.

The concept of agency can also facilitate the derivation of testable predictions from a model, once it has been formulated. This is perhaps most common when the derivation is done informally, such as through verbal reasoning about how organisms would best realize their agenda in particular situations. But the agential view can also be pursued formally, with explicit mathematical framing of a maximization problem. Even when formal derivation of mathematical results is done without recourse to notions of agency, these same results are very often translated into

agential language for the purpose of conceptualization and understanding, which are also important parts of the process of making testable predictions.

Crucially, no aspect of the optimality approach commits biologists to the view that any particular trait is an adaptation or that it is at its optimum. Just as evolutionary biologists make powerful use of Maynard Smith and Price's (1973) "evolutionarily stable strategy" concept for the purpose of making comparative predictions whilst fully acknowledging that no real-world biological population is ever truly at equilibrium, we also make powerful use of the optimality approach for deriving clear, testable predictions about organismal traits whilst fully acknowledging that no trait is ever truly at its optimum (Gardner 2009).

The optimality approach does not commit evolutionary biologists to the belief that all traits are perfectly optimized adaptations, but it does require that we specify *who* the agent is and *what* is their goal. For example, insofar as the interests of an individual organism differ from those of her wider social group, then it makes a substantive difference whether we take an individual-level versus a group-level optimality approach. Happily, such decisions need not be made in an ad hoc way, but can be informed by evolutionary theory.

Okasha discusses Fisher's (1930) "fundamental theorem" of natural selection, and warns that it provides no firm basis for supposing that cumulative evolutionary change across multiple generations results in perfect adaptation. I believe this misses the point. The central importance of the fundamental theorem to evolutionary biology is not that it predicts precise adaptation but rather that it reveals the identity of the adaptive agent, i.e. the individual organism, and pinpoints her agenda, i.e. maximization of her fitness (Gardner 2017).

The fundamental theorem properly applies only to the special case in which there is no social interaction between genetic relatives (Special Darwinism), as Fisher made clear (Gardner 2017). However, a modified version of the fundamental theorem can be obtained for the more general situation in which social interaction may occur between genetic relatives (General Darwinism), and this continues to identify the individual organism as the adaptive agent but clarifies that her goal is actually to maximize her *inclusive fitness* (Hamilton 1964; Gardner 2017). Again, the importance of this result is not that it predicts perfect social adaptation, but rather that it identifies the adaptive agent and spells out her agenda.

Multilevel selection provides an alternative approach to understanding social adaptation (Price 1972; Hamilton 1975). This yields a further version of the fundamental theorem that highlights how the social group, too, can be regarded as an agent with a goal of fitness maximization, but only if within-group selection is absent or negligible—as, for example, in the case of clonal groups (Gardner 2015a, 2017). Once again, this fundamental theorem of multilevel selection says nothing about the precision of adaptation. Instead, its relevance is that it establishes when it is valid for the optimality approach to focus on group interest.

Okasha mischaracterizes my position on group-level adaptation when he says that I have suggested that all traits exhibited by clonal groups *must* be understood as group-level adaptations. In reality, my view is simply that they *may* be conceptualized in this way. The individuals who make up a clonal group may, of course, also be viewed as fitness-maximizing agents in their own right. For example, a white blood

cell that engulfs a bacterium can be viewed either as an inclusive-fitness-maximizing cell or as one facet of a whole-organism response to infection (Gardner 2015b).

In fact, I feel it is Okasha who is too prescriptive as to the assignment of adaptive rationale in the context of clonal groups. He makes the classic distinction between a fleet herd and a herd of fleet individuals to argue that, unless there is an explicitly social aspect to a trait exhibited by a clonal group, then the trait *must* be viewed as an individual-level adaptation and *cannot* be viewed as a group-level adaptation. I believe this disagreement underlines the crucial difference between his philosophical outlook and my own scientific approach. My view is that, as individual and group optima exactly coincide under clonality, both individual and group optimality approaches yield exactly the same testable prediction, and so from a scientific perspective, both are equally valid interpretations (Gardner 2015b). Again, my work on group-level optimality is not motivated by a need to characterize this or that trait as a group-level adaptation, but rather to establish when a group can validly be considered an adaptive agent in its own right for the purpose of making scientific predictions about all of its traits.

I also find myself disagreeing with Okasha over the gene's eye view of evolution. Applying the fundamental theorem approach at this lower level of biological organization, I feel that the gene is most naturally viewed as a miniature organism that must, on account of its social interactions with other genes that may bear the same allele, be regarded as having an agenda of inclusive-fitness, rather than personal-fitness, maximization (Gardner and Welch 2011). That is, the gene is a concrete token rather than an allelic type, and its agenda in particular instances may be altruistic rather than always selfish. This view clearly deviates from Dawkins' (1976) selfish-allele approach, which Okasha endorses, but it appears more consistent with how evolutionary biologists actually speak of the interests of genes—particularly in relation to the intragenomic conflicts that Okasha believes most strongly motivate taking the gene's eye view at all (Gardner and Úbeda 2017).

For example, Haig's (2002) insight that an "offspring's paternally derived genes have different interests from the offspring's maternally derived genes" only makes sense if we understand these genic agents to be physical scraps of DNA rather than allelic types, as it is only the physical scrap of DNA that can be described as having, say, a maternal origin. The allele, potentially present in multiple bodies simultaneously, may in some instances be derived from its carrier's mother and in others from its carrier's father. Accordingly, this intragenomic conflict *must* be between scraps of DNA, rather than between alleles (Gardner and Úbeda 2017). More generally, this question of the gene's identity seems like philosophically, as well as scientifically, rich territory, so I am a little disappointed that Okasha has chosen to pass over the topic of genic agents so quickly and without probing these issues.

In conclusion, whilst there is much that I like about Okasha's book, I regret that there is not more science. Okasha says that the ultimate justification for the agential view of evolution is empirical rather than theoretical, and I suppose he considers the empirical to be outwith his purview. In contrast, I believe that the ultimate justification for the agential view is *scientific*. Scientific research programmes are judged according to how well they facilitate the formulation of new hypotheses and the derivation of testable predictions, and ultimately by how well they advance scientific

understanding through the retention and refinement of certain hypotheses and the rejection of others. This requires a tight interplay of the theoretical and the empirical—both are equally important. Exploring how this process works in particular disciplines, such as evolutionary biology, and evaluating the conceptual tools of the trade within this context, is certainly within the remit of philosophy of science.

References

Dawkins, R. 1976. *The selfish gene*. Oxford: Oxford University Press.

Fisher, R.A. 1930. *The genetical theory of natural selection*. Oxford: Clarendon Press.

Gardner, A. 2009. Adaptation as organism design. *Biology Letters* 5: 861–864.

Gardner, A. 2015a. The genetical theory of multilevel selection. *Journal of Evolutionary Biology* 28: 305–319.

Gardner, A. 2015b. More on the genetical theory of multilevel selection. *Journal of Evolutionary Biology* 28: 1747–1751.

Gardner, A. 2017. The purpose of adaptation. *Interface Focus* 7: 20170005.

Gardner, A., and F. Úbeda. 2017. The meaning of intragenomic conflict. *Nature Ecology and Evolution* 1: 1807–1815.

Gardner, A., and J.J. Welch. 2011. A formal theory of the selfish gene. *Journal of Evolutionary Biology* 24: 1801–1813.

Haig, D. 2002. *Genomic imprinting and kinship*. New Brunswick: Rutgers University Press.

Hamilton, W.D. 1964. The genetical evolution of social behaviour. *Journal of Theoretical Biology* 7: 1–52.

Hamilton, W.D. 1975. Innate social aptitudes of man: An approach from evolutionary genetics. In *Biosocial anthropology*, ed. R. Fox, 133–155. New York: Wiley.

Harvey, P.H., and M.D. Pagel. 1991. *The comparative method in evolutionary biology*. Oxford: Oxford University Press.

Maynard Smith, J., and G.R. Price. 1973. The logic of animal conflict. *Nature* 246: 15–18.

Okasha, S. 2018. *Agents and goals in evolution*. Oxford: Oxford University Press.

Parker, G.A., and J. Maynard Smith. 1990. Optimality theory in evolutionary biology. *Nature* 348: 27–33.

Price, G.R. 1972. Extension of covariance selection mathematics. *Annals of Human Genetics* 35: 485–490.