© 2014 John Wiley & Sons Ltd
Ratio (new series) XXVIII 1 March 2015 0034-0006

LEVELS OF SELECTION ARE ARTEFACTS OF DIFFERENT FITNESS TEMPORAL MEASURES

Pierrick Bourrat

Abstract

In this paper I argue against the claim, recently put forward by some philosophers of biology and evolutionary biologists, that there can be two or more ontologically distinct levels of selection. I show by comparing the fitness of individuals with that of collectives of individuals in the same environment and over the same period of time – as required to decide if one or more levels of selection is acting in a population – that the selection of collectives is a by-product of selection at the individual level; thus, talking about two or more levels of selection represents merely a different perspective on one and the same process.[1]

1. Introduction and Background

Although the idea of group selection had been largely rejected by the end of the 1960s, it has recently undergone a revival (e.g. Okasha, 2006; Sober & Wilson, 1998). Its proponents argue that, in principle, an ontologically distinct process of selection can occur at each level of organisation, and that there are some real instances of this in biology. Following a classical definition, I define a process of selection at one given level in a population as the elimination of one type of entity in favour of at least one other due to a *systematic* difference in physical properties between the different types forming this population. Such systematic differences in physical properties are what biologists and philosophers of biology call differences in fitness.

Two important areas of evolutionary biology in which the multilevel selection idea has been applied are species selection

---

[1] I am particularly thankful to Kristie Miller and David Braddon-Mitchell who both encouraged me writing this paper. I am also thankful to Paul Griffiths, Adam Hochman, Robyn Kath and an anonymous reviewer Gladys Kostyrka, for their comments on an earlier version of the manuscript. I thank particularly Robyn Kath who had a thorough final check of the manuscript. This research was supported under Australian Research Council's Discovery Projects funding scheme DP0878650 and The International Postgraduate Research Scholarships from the University of Sydney.

(Lewontin, 1970; Okasha, 2006) and evolutionary transitions in individuality (see Calcott & Sterelny, 2011 for a survey of the most recent work on this subject; see also Godfrey-Smith, 2009; Okasha, 2006). In this paper I show that the claim that there can be one level of selection acting at each level of organisation, although becoming increasingly popular, is unwarranted, and is an artefact of different temporal measures of fitness. Bouchard (2008, 2011) recently argued that comparisons of fitness over absolute time is a superior approach in evolutionary theory than comparisons over one or more generations. The main reason advanced by Bouchard is that some organisms do not reproduce, yet it would be odd to claim that they have nil fitness. Here I provide another argument that shows why absolute time is better suited than generations to the measurement and comparison of fitness. Although I find multilevel selection models perfectly reasonable and useful tools, my aim is to show that their ontological interpretation is problematic.

2. Fitness and Meaningful Comparisons of Fitness

The concept of fitness, although fundamental to our understanding of the process of natural selection, is not a unified concept and many problems surrounding its interpretation remain (Abrams, 2012; Rosenberg & Bouchard, 2008). In philosophy of biology the most popular account of fitness is known as the propensity interpretation of fitness (Beatty & Finsen, 1989; Brandon, 1978; Mills & Beatty, 1979; Pence & Ramsey, 2013; Sober, 1984, 2001). Proponents of this account hold that fitness is the disposition of an entity to produce a certain number of offspring with a probability associated to each possible outcome. Disposition is understood here in the same way as the disposition of a piece of glass to break. For example, the propensity of an entity to produce offspring might be a 50% chance to produce no offspring, a 25% chance to produce one offspring, and a 25% chance to produce two offspring. Based on their propensity, each entity under the propensity interpretation of fitness has an *expected* number of offspring (in our case 0.75), which represents perfectly, with a single number, the fitness of this entity. Some entities have a disposition to produce more offspring overall (a higher expected number of offspring) in the same way that some pieces of glass have a disposition to break more easily.

© 2014 John Wiley & Sons Ltd

Entities can be thought of as *biological organisms*, but there is a long tradition of expanding the concept of natural selection beyond biological organisms. Under a "recipe" formulation of evolution by natural selection, any sort of entities forming populations and exhibiting heritable variation in fitness will evolve by natural selection (Godfrey-Smith, 2009; Lewontin, 1970). The propensity interpretation of fitness was initially proposed to solve the tautology problem of fitness: if fitness is defined in terms of offspring produced by an individual, fitness simply restates the number of offspring it produces without causally explaining why it produced this number of offspring rather than another one. With a particular reproductive output causally traceable to a propensity, the fitness of an individual becomes non tautologous (Pence & Ramsey, 2013).

As popular as it is, this account suffers from major counterexamples (see Beatty & Finsen, 1989; Godfrey-Smith, 2009; Pence & Ramsey, 2013; Sober, 2001). But maybe the most important problem for the propensity interpretations of fitness, although rarely mentioned, is that among the different interpretations of probabilities, propensities interpretations are very controversial (Eagle, 2004; Hájek, 2003). In fact, it is not clear what sort of properties propensities are. This has led some, such as Byerly and Michod (1991), to argue that fitness, when it is understood as a propensity, is empirically empty because it fails to explicate what the causal connections are, between reproductive outputs and the functional relations an entity has with its environment. Yet ecology is a critical component of evolutionary biology. Put another way, there is no way to obtain empirically the propensity of a given entity to produce offspring: propensities must be known *a priori*. In spite of these problems, because it is the most popular account of fitness in the philosophy of biology and because this has no impact on the problems tackled in this article, I will follow the propensity interpretation of fitness. I define the fitness of an entity as its expected reproductive output and, in cases in which the entity does not reproduce, as its expected time to extinction or expected growth rate or more generally, following Bouchard (2008, 2011), as persistence through time.

But whatever definition of fitness one decides to use, it is clear that the fitness of an entity can only be estimated empirically by measuring *actual* or *realized* reproductive outputs, rates of extinction or growth rates. I call these measures *fitness proxies*. As I noted in the introduction, natural selection, of which

© 2014 John Wiley & Sons Ltd

multilevel-selection is a complex form (one process of natural selection occurring at each level of organisation), is often characterized as resulting from a difference in fitness. Yet, not any sort of difference in fitness proxies is a good estimate of differences in fitness in the sense relevant to natural selection.

To understand why, one can start by observing that no biologist would compare the reproductive output of a mouse with that of an elephant and claim, based on the fact that the average reproductive output of a mouse is higher than that of an elephant, that the mouse is fitter than the elephant. An average reproductive output of ten after one generation for a mouse cannot be compared to an average reproductive output of one after one generation for an elephant, when one intends to make claims about natural selection. This is because the environments of mice and elephants are very different and one generation for an elephant is a much longer time than one generation for a mouse.

The importance of the environment in comparing organisms' proxies of fitness is not limited to the interspecific case.[2] In fact, it would be a mistake to compare the reproductive output of elephants of two types, if the two types are found, on average, in different environmental conditions. The difference in environmental conditions undergone by each type might be the reason why they have, on average, different reproductive outputs – not that one had a propensity to produce more offspring. To claim that one type of elephant is fitter than the other based on different average reproductive outputs only, one would need to compare the reproductive output of the two types in every possible environmental condition. This would require an unlimited population size of elephants because there are an unlimited number of possible environmental conditions. This solution is unmanageable empirically. Another approach, more manageable, is to compare differences in reproductive outputs between the two types under the same environmental conditions, so that the effects of the environment on the reproductive outputs of the two types cancel each other out. Then, only a physical difference between the two types can explain the difference in reproductive output, which is thus attributable to a difference in fitness. Although this solution is empirically more manageable, it is also at greater risk of being wrong if one forgets a confounding variable in the comparison.

---

[2] This problem is explored at length in Brandon (1990).

© 2014 John Wiley & Sons Ltd

To sum-up so far: to avoid confounding variables when comparing fitness proxies (to infer which one of two or more entities is fitter and will be selected) one must ensure that the entities compared are in the same environment. This requires that the fitness proxies used are measured over the same period of time. If the fitness proxies of two entities are measured over different periods of time, then the two entities have undergone different factors that could have consequences on reproductive output (or more generally, on any proxy of fitness) and thus the measurements are made in different environments. An environment should thus be defined both spatially and temporally.

Why are those considerations relevant to the idea of multiple levels of selection? When evaluating whether a population is undergoing a different process of selection at several levels of organisation, one would ideally want to compare fitnesses at those different levels *directly*. Yet, in such a case, as in any other case, only fitness proxies (e.g. reproductive outputs of organisms, rates of extinctions of species, etc.) are available to the observer. This means that one must compare those fitness proxies in the same environment, and thus over the same period of time, so as to register only differences in fitness and not differences in the environment.

The comparison involved in testing whether there is more than one process of selection acting at two[3] different levels in a population is threefold and I will call it ***C***. The first step to make ***C*** is to calculate a difference in fitness proxy at the lower level between at least two types of entities under the same environment, *E*, and thus over the same period of time, *t*. Second, the same kind of comparison must be made at the higher level of organisation between the higher level entities that supervene on the lower ones in *E*, again over *t*. Finally, the sign of the two differences calculated must be compared. Only if the two differences are of opposite sign one can conclude with certainty[4] that there is one level of selection at each level. Basically, the third step is equivalent to a clean intervention on the level of organisation, using Woodward's (2003) terminology. Only if the level of organisation is the sole difference maker in the direction of selection at different levels of organisations, can one conclude that each level is a genuine level of selection. In the following section, I present the classical argu-

---

[3] But the same reasoning can be applied to more than two levels.

[4] If the signs are the same, no conclusion can be drawn since there could be two different selection processes going in the same direction.

© 2014 John Wiley & Sons Ltd

ment that led some authors to consider that there are genuine levels of selection. In section four, I show that this argument is flawed for it is not in line with the requirements of ***C***. From there, I draw the consequences for the claim that there can be ontologically distinct levels of selection, which is based on this flawed argument.

3. Rationale behind the idea of selection processes at multiple levels

Imagine two types of entities (e.g. individual organisms) that we will call A and B. Type A pays a cost to reproduce, which is then translated into having less offspring than type B which does not pay this cost. Assuming that A and B are in the same environment, *E*, type B is thus fitter than type A. Now imagine a case in which members of type A gather together to form collectives of type A and similarly members of type B form collectives of type B. Suppose that once these two types of collectives are put together in *E*, they beget new collectives of their type, but collectives A are more fecund than collectives B. It seems here that collectives A are fitter than collectives B. How is it possible, at the same time, that individuals B are fitter than individuals A but that collectives A are fitter than collectives B? The answer given by the proponents of the idea of multiple levels of selections is that there is one distinct process of selection operating at each level. One process, they argue, pushes in one direction at the individual level, favouring individuals B over individuals A. Another, at the collective level, pushes in the other direction, favouring collective A over collective B.

Okasha (2006)[5] uses this type of reasoning when discussing two important topics of evolutionary theory, namely species selection and evolutionary transitions in individuality (see also Michod, 2005). To make the reasoning more concrete, let us imagine that individuals of type A are organisms reproducing sexually and individuals of type B are organisms reproducing asexually.[6] Let us also imagine that collectives A are sexual species and collectives B are asexual species belonging to the same reference class of

---

[5] Okasha borrows his reasoning from Vbra (1989).

[6] The same reasoning could be made while using the notion of evolutionary transition in individuality whereby selfish individuals are fitter than altruistic individuals, but collectives of altruistic individuals are fitter than collectives of selfish individuals.

© 2014 John Wiley & Sons Ltd

species as A (i.e. species with a similar ecology). We know from biology that, everything else being equal, sexual organisms pay an evolutionary cost when compared to asexual organisms. The so-called 'twofold cost of sex' (Maynard Smith, 1978) arises because only half of sexual organisms can reproduce *and* must produce offspring of both sexes. Under a constant environment asexual organisms are fitter than sexual organisms. On the other hand, there is a general tendency for sexual species to have a rate of extinction lower than asexual species (Gouyon, 1999; Maynard Smith, 1986; Nunney, 1989), for asexual species are often considered to be evolutionary dead ends. Put together under the same regime, sexual species do better than asexual ones: they persist longer. Thus, it is tempting to consider that asexual organisms are fitter than sexual organisms, that sexual species are fitter than asexual species, and take this as the demonstration that there are two distinct levels of selection operating at two levels of organisation.

4. What is going wrong?

All would be well with this reasoning if by "multiple levels of selection" one was referring to a tool or methodology for partitioning a complex selection process occurring at the individual level, recognizing that it represents only a heuristic. Talk of multiple levels of selection would simply amount to a convenient way to partition off this unitary, yet hard to visualize, process into components. But that is not what is meant by the proponents of this idea: they are ontologically committed to the idea of distinct processes acting at different levels. One example where this commitment is apparent is in the exchange between Waters (2011) and Okasha (2011) in which Waters, reviewing the work of Okasha (2006), proposes that the notion of levels of selection in the context of species selection is merely a conventional one. Okasha (re)asserts that it is an ontological one. Yet, as I demonstrate below, although the heuristic notion of multiple levels of selection is sound, the ontological one is not. I show that the classical reasoning behind the idea of multiple (ontological) levels of selection does not conform to the requirements of ***C*** as presented in the previous section.

I remind the reader that conformity to ***C*** is needed for making comparisons of fitness proxies between levels that will allow us to

© 2014 John Wiley & Sons Ltd

decide whether selection operates at more than one level.[7] To begin with, note that one difference between the measures of fitness proxies between organisms and species is that they are usually made over vastly different periods of time. The reproductive output of an organism is usually much shorter than the time to extinction of a species.[8] Thus, to be precise, asexual organisms have more offspring than sexual organisms *over a short period of time* (that might be one organism generation, but is not necessary), while sexual species go extinct less often than asexual species *over a long period of time*. This difference is in itself not problematic since having measures of a given phenomenon at different scales (whether spatial or temporal) is usually good practice in science.

However, this can be problematic for our aims if the variability of the environment is positively correlated with time. Remember that by applying ***C*** we want to know whether one process of selection pushes in one direction at the organism level and another process pushes in another direction at the species level. But with proxies of fitness over different periods of time for each level (instead of fitness), two alternative hypotheses can explain the different directions of evolutionary change observed at the two levels. As will become apparent, each hypothesis makes different assumptions about the variability of the environment, *E*, over time. The first hypothesis (H1) is that environmental conditions are more variable over long periods of time than shorter ones and that asexuality (or in the formal case type B) is fitter in stable (or quasi-stable) conditions and sexuality (type A) is fitter when the environmental conditions change.

Under H1, it is correct to assert that asexual organisms are fitter than sexual organisms – but only assuming a constant environment (*Ec*). Similarly, it is correct to assert that sexual species are fitter than asexual species – but only when the conditions are unstable (*Eu*). The fact that we observe a change of direction in evolutionary change between the two levels merely reflects a change in temporal scale, which itself is a proxy for environmental changes. Thus under H1, the level of organisation/selection does not play any causal role in the differences in fitness observed; there is only one process of selection, pushing in one direction

---

[7] In the rest of the paper, I use the asexual/sexual reproduction example for its appeal when compared to the formal case of types A and B, but the same demonstration could be done with any isomorphic case.

[8] The same argument holds if the fitness of a species is understood as a number of offspring species.

© 2014 John Wiley & Sons Ltd

when the environment is stable and in another when the environment is unstable. Had the fitness of *organisms* been measured under *Eu,* then the fitness of *sexual organisms* would have been *higher* than the fitness of asexual organisms. Similarly, had the fitness of *species* been measured under *Ec,* the fitness of *asexual species* would have been *higher* than the fitness of sexual species.

The second hypothesis (H2) is the one proposed by the ontologically committed proponents of multiple levels of selection, i.e. that there are two ontologically distinct levels of selection, and that even if measured over the same period of time we would observe a different direction of selection at each level. This would lead to the conclusion that there is one ontologically distinct level of selection operating at each level, thus two (or more) processes of selection. Under H2, in spite of the differences in time scale for the measures of fitness at different levels, *C* can still be applied because this difference does lead to overall differences in the environment, *E,* at the two levels. Variability in *E* is not correlated with measures of fitness over different time scales.

There are two reasons why H1 should be preferred over H2. First, one reasonable assumption to make is that the shorter the period of time over which fitness is measured the more probable *E* will remain stable over this period of time. If one accepts this assumption (which concurs with our scientific knowledge about past environments), there is always a possible confounding variable that can explain why selection goes in opposite directions at two levels when measuring fitness over different periods of time, whether at one or several levels. As a consequence, it makes H2 weaker than H1. Without a "clean" control or intervention over the level of organisation (i.e. one that does not modify any other variable), not only can no conclusion be drawn about selection at these levels, but H2 is a less parsimonious hypothesis than H1.

Second, noting that species, and collectives more generally, are made up of organisms or lower level entities (collectives supervene on lower level entities), if one keeps the measures of fitness over the same period of time, it is hard to see how two processes of selection could go in opposite directions when one and the same population is considered at each level. Arguing that two such processes exist would require the invocation of some emergent properties or emergent levels of selection (whatever that means) in a strong sense. For H2 to be acceptable, a proponent of this hypothesis should explain what benefit would be gained by inflating our ontology when an alternative hypothesis (namely

© 2014 John Wiley & Sons Ltd

H1) is available, which does not require any ontic inflation (i.e. postulates only one level of selection), predicts the same results and recognizes the pragmatic value of carving a selection process both spatially and temporally.

5. Conclusion

In this paper I have shown that the idea that there can be distinct processes of selection acting at each level of organisation is not as straightforward as its proponents claim it to be. There are good reasons to think that what looks like two distinct levels of selection is in fact one and the same process of selection operating over different environmental conditions, as a result of proxies of fitness being measured over different time scales. Fitness does not emerge at new levels of organisation; only our appreciation of it changes when we straddle the levels of organisation. This conclusion does not render the idea of multiple levels of selection useless, since it might be useful to consider selection and fitness over different periods of time (Abrams 2009, Sober 2001, Mills and Beatty 1979). Shifting from one level to another might be a convenient proxy for doing that. There are many ways to model evolutionary changes. The multilevel approach is one of them and it is a useful one. But the view that there are ontologically distinct processes of selection is merely a mirage.

School of Philosophical and Historical Inquiry
Department of Philosophy
The University of Sydney,
Main Quadrangle A14
Sydney, NSW 2006, Australia
p.bourrat@gmail.com

Works cited

Abrams, M. (2009). The Unity of Fitness. *Philosophy of Science*, 76, no. 5.
—— (2012). Measured, Modeled, and Causal Conceptions of Fitness. *Frontiers in Genetics*, 3, 196.
Beatty, J. H. and Finsen, S. K. (1989). Rethinking the propensity interpretation: A peek inside Pandora's box. In M. Ruse (Ed.), *What the Philosophy of Biology Is: Essays Dedicated to David Hull*. Dordrecht: Kluwer Publishers.
Bouchard, F. (2008). Causal processes, fitness, and the differential persistence of lineages. *Philosophy of Science*, 75(5), 560–570.

© 2014 John Wiley & Sons Ltd

—— (2011). Darwinism without populations: a more inclusive understanding of the “Survival of the Fittest”. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 42(1), 106–114.

Brandon, R. N. (1978). Adaptation and evolutionary theory. *Studies In History and Philosophy of Science Part A*, 9(3), 181–206.

—— (1990). *Adaptation and environment*. Princeton: Princeton University Press.

Byerly, H. C. and Michod, R. E. (1991). Fitness and evolutionary explanation. *Biology and Philosophy*, 6(1), 45–53.

Calcott, B. and Sterelny, K. (2011). *The major transitions in evolution revisited*. Cambridge: The MIT Press.

Eagle, A. (2004). Twenty-one arguments against propensity analyses of probability. *Erkenntnis*, 60(3), 371–416.

Godfrey-Smith, P. (2009). *Darwinian populations and natural selection*. USA: Oxford University Press.

Gouyon, P. (1999). Sex: a pluralist approach includes species selection.(One step beyond and it’s good.). *Journal of Evolutionary Biology*, 12, 1029–1030.

Hájek, A. (2003). Interpretations of probability. *The Stanford Encyclopedia of Philosophy*.

Lewontin, R. C. (1970). The units of selection. *Annual Review of Ecology and Systematics*, 1(1), 1–18.

Maynard Smith, J. (1978). *The evolution of sex*. Cambridge: Cambridge University Press.

—— (1986). Evolution: contemplating life without sex. *Nature*, 324, 300–301.

Michod, R. E. (2005). On the transfer of fitness from the cell to the multicellular organism. *Biology and Philosophy*, 20(5), 967–987.

Mills, S. K. and Beatty, J. H. (1979). The propensity interpretation of fitness. *Philosophy of Science*, 263–286.

Nunney, L. (1989). The maintenance of sex by group selection. *Evolution*, 43(2), 245–257.

Okasha, S. (2006). *Evolution and the Levels of Selection*. USA: Oxford University Press.

—— (2011). Reply to Sober and Waters. *Philosophy and Phenomenological Research*, 82(1), 241–248.

Pence, C. H. and Ramsey, G. (2013). A new foundation for the propensity interpretation of fitness. *The British Journal for the Philosophy of Science*, 64(4), 851–881.

Rosenberg, A. and Bouchard, F. (2008). Fitness. *Stanford encyclopedia of philosophy*.

Sober, E. (1984). *The nature of selection*. MA: MIT Press Cambridge.

—— (2001). *The two faces of fitness Thinking about evolution: historical, philosophical, and political perspectives*. In R. S. Singh, C. B. Krimbas, D. B. Paul & J. Beatty (Eds.), Cambridge: Cambridge University Press.

Sober, E. and Wilson, D. S. (1998). *Unto Others: The evolution and psychology of unselfish behavior*: Cambridge: Harvard University Press.

Vbra, E. S. (1989). Levels of selection and sorting with special reference to the species level. *Oxford Surveys of Evolutionary Biology*, 6, 111–168.

Waters, K. C. (2011). Okasha’s Unintended Argument for Toolbox Theorizing. *Philosophy and Phenomenological Research*, 82(1), 232–240.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. USA: Oxford University Press.

© 2014 John Wiley & Sons Ltd