The Causal Structure of Evolutionary Theory

Grant Ramsey

Australasian
Association of
Philosophy

Routledge
Taylor & Francis Group

The Causal Structure of Evolutionary Theory

Grant Ramsey

Katholieke Universiteit Leuven

ABSTRACT

One contentious debate in the philosophy of biology is that between the statisticalists and causalists. The former understand core evolutionary concepts like fitness and selection to be mere statistical summaries of underlying causal processes. In this view, evolutionary changes cannot be causally explained by selection or fitness. The causalist side, on the other hand, holds that populations can change in response to selection—one can cite fitness differences or driftability in causal explanations of evolutionary change. But, on the causalist side, it is often not clear how, precisely, one should understand these causes. Thus, much more could be said about what sort of causes fitness and driftability are. In this paper, I borrow Dretske's distinction between structuring and triggering causes and I suggest that fitness and driftability are structuring causes of evolution.

1. Introduction

One of the central debates in the philosophy of biology concerns the causal interpretation of the theory of evolution by natural selection. This debate has become polarized into two camps, the causalists and the statisticalists. The causalists, by and large, argue that changes in a population can be causally due to fitness differences or to selection: selection can cause changes (or stasis) in the frequencies of organism types in a population, or can shift (or stabilize) the mean value of quantitative traits, such as height [Sober 1984; Brandon 1990; Bouchard and Rosenberg 2004; Millstein 2006; Brandon and Ramsey 2007; Ramsey 2013a, 2013c]. Drift, too, has been given a causal construal. Drift (often labeled 'genetic drift') is standardly defined by biologists as a change in the frequency of gene variants (alleles) in a population due to random sampling. Despite the simplicity of the biological definition of drift, there has been considerable debate concerning the fundamental nature of drift and how it can be distinguished from selection. Furthermore, both biologists and philosophers at times talk of evolutionary change being due to drift, where drift is considered not to be a mere outcome, but instead that which causally explains the outcome [Millstein 2002; Gildenhuys 2009; Ramsey 2013b].

In contrast with the causalists, the statisticalists argue that causal aspirations for fitness, selection, and drift cannot be realized. Fitness, selection, and drift are, they argue, outcomes, not causes. Population changes do have causes, but these are at the level of individual births, deaths, and other occurrences, not at the level of fitness differences

[Matthen and Ariew 2002; Walsh, Lewens, and Ariew 2002; Pigliucci and Kaplan 2006; Walsh 2007, 2010; Ariew and Ernst 2009].

Although the causalists have offered a compelling defence of their position, their arguments have not been entirely satisfying. For one thing, there is a lack of unanimity on the causalist side as to which of the central evolutionary concepts can causally explain evolutionary outcomes. For example, some have suggested that fitness differences can help to causally explain evolutionary outcomes [Ramsey 2013a], while others have argued that it is selection, not fitness, that is causally effective [Sober 1984]. And some have argued that it is not selection in general that is causal, but instead that it is 'selection-for' where the causal rubber meets the road [ibid.]. Causalists are also divided on the concept of drift, debating whether drift is a causal process [Millstein 2002], a mere product [Brandon 2005], or whether it can play roles both as cause and as effect [Ramsey 2013b].

Which (if any) of these causalist positions is correct? This paper is an attempt to describe a unified account of the causal structure of evolution by natural selection. In describing the account, I borrow a distinction from Dretske [1988] between structuring and triggering causes to help us to understand how natural selection, fitness, and drift-ability can cause evolutionary change.

2. Structuring and Triggering Causes

Dretske [ibid.] introduced a distinction between structuring and triggering causes. This distinction is part of a family of distinctions, some of which preceded Dretske. For example, Mackie [1965] distinguishes predisposing causes (causal conditions that set the stage for an event to occur) from triggering causes (causes that trigger the event's occurrence), and the language of precipitating versus predisposing causes is common in the medical literature (smoking can predispose you to have cancer without the cancer actually developing—that is, without it being precipitated). And in the causal explanation literature there are parallel distinctions, like that between process and program explanations [Jackson and Pettit 1992]. Program explanations cite causes that ensure ('program for') particular outcomes, whereas process explanations identify the actual causal chain leading to the outcome. Finally, Sterelny's [1996] distinction between actual-sequence and robust-process explanations parallels that from Jackson and Pettit: sometimes we explain via actual sequences/processes, while at other times we explain by pointing to the robustness or inevitability of the outcome.

I will use Dretske's [2004] formulation here, since it has a less teleological connotation than does talk of a program, since outcomes need not be robust, and because the concept of structure nicely maps onto the ideas below concerning the structure of life histories. In Dretske's rendering, the distinction is that structuring causes set up the structure of a causal system, while triggering causes trigger the system to produce its effect. He uses an example of a terrorist and a general to illustrate this distinction [ibid.: 169]:

> A terrorist plants a bomb in the general's car. The bomb sits there for days until the general gets in his car and turns the key to start the engine. The bomb is detonated (triggered by turning the key in the ignition) and the general is killed. Who killed him? The terrorist, of course. How? By planting a bomb in his car. Although the general's own action (turning on the engine) was the triggering cause, the terrorist's action, wiring the bomb to the ignition, is the structuring cause, and it will surely be the terrorist's action, something that happened a week ago, that will be
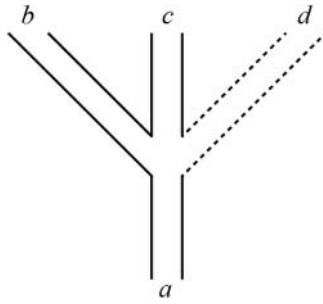
Figure 1. The terrorist structured the situation at *a* so that the car will either explode (*b*) or create smoke (*c*). The structuring cause does not allow for the horn to be triggered (*d*).

> singled out, in both legal and moral inquiries, as the cause of the explosion that resulted in the general's death.

This is a simple example with one setup and a single outcome (the car explosion). But more complex examples are possible. It could be that the terrorist did not want to blow up the general's family, so she put a pressure sensor under the back seat that would make the key trigger only a secondary smoke bomb intended to warn and to terrorize him and his family. A single structuring cause (setting up the car with the pair of bombs, the pressure sensor, and the key switch) has set the world to have two possible outcomes (assuming, of course, that the general will definitely turn the key).

If the probability that the general is with his family when he next drives his car has a value between one and zero, then there are intermediate probabilities for each outcome (fig. 1). Other possible outcomes, such as the key's triggering a loud horn to sound, are not possible given the structuring cause (see *d* in fig. 1). This probabilistic branching structure is important because, I will argue, this is the kind of causal structure that is realized in evolution by natural selection. But before we can get there, let's consider a natural structuring cause.

3. Of Fruit and Rivers

The case of the general involved two agents—a general and a terrorist—and, in the case of the smoke bomb, two probability-weighted outcomes. This example might seem to imply that structuring causes require agency, or at the very least some sort of planning, anticipation, or conceptualization. But this isn't the case. Now consider a setup in which there are no agents involved, but where outcomes nevertheless can be understood in terms of their structuring and triggering causes.

Consider a river that forks into two branches, where three-quarters of the water passes through the left branch, while one-quarter passes through the right. Now consider a fruit floating down the river. The fruit falls into the river from an overhanging branch at point *a* (in fig. 2) and ends up either at point *b* or point *c* with a 0.75 probability of the former and a 0.25 probability of the latter.[1] Point *d* is on dry land and the

---

[1] In order for these probabilities to obtain, assumptions must be made about the nature of the water flow, such as that it is turbulent instead of laminar, or that random external influences (like the wind) ensure that the probability density is uniform across the cross section of the river at the place where it forks.
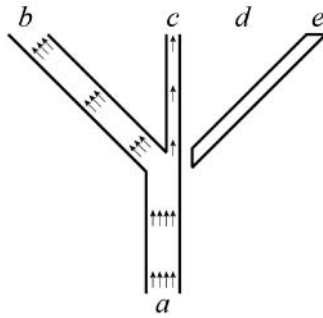
Figure 2. Fruit falls into a river at *a*, has a 0.75 probability of ending up at *b*, a 0.25 probability of ending up at *c*, and a 0 probability of ending up at *d* or *e*.

river used to have a second branch ending at point *e*, but the entrance to this branch has become silted in and is now just a small, stagnant lake.

It turns out that the fruit under consideration ends up at point *b*. What are the structuring and triggering causes in this example and what causally explains this outcome? Consider first the triggering cause. Like the turning of the key in the case of the general, here we have the fruit falling into the stream at point *a*: What triggered the general's death? He turned the key. What triggered the fruit to end up at *b*? It fell in the river at *a*.[2]

What about a structuring cause in such a case? While the triggering cause of the fruit's destiny is its plunge into the river, the structuring cause of the fruit ending up at *b* is the erosion, deposition, and other forces having set the river's shape. Given this shape and the flowing water, we have a causal explanation for why the fruit ended up at *b* instead of, say, *d* or *e*. The riverbed morphology prevents the fruit from ending up on dry land, *d*, or in the isolated lake, *e*. Considering the question of why the fruit ended up at *b* instead of *c*, the structuring cause results in intermediate probabilities for each outcome, but does not say with certainty which one will occur for each fruit in question.

This situation is like the case of the terrorist and the general. The structuring cause can be used to explain why the general either died (*b*), or was subject to a smoke bomb (*c*). But it does not explain why one of these rather than the other occurred. This explanation is due to the triggering cause: either the general alone triggered the ignition switch, or the family triggered the pressure-sensitive switch and the general triggered the ignition switch. Which one of these occurred could be due to a host of external variables. If, for example, it is unusually cold, the general may go to start the car early in order for it to warm up before his family arrives. And because which effect is triggered depends on such variables, the structuring cause is not determinative of these outcomes. It does, however, explain why a loud horn does not sound when the general turns the key (*d*). The terrorist could have chosen this as a scare tactic instead of choosing the smoke, but the fact that she decided on the smoke bomb and wired the car

---

[2] The reason why this is 'fell into the river at *a*' instead of 'fell into particular place *p* in the river at *a*' is that, depending on what happens between *a* and the split leading to *b* and to *c* (and what the natures are of the probabilities referred to in note 1), the specific details of where the fruit landed in the river may, or may not, carry information about its chances of ending up in *b* or in *c*.

accordingly allows one to use the structuring cause to explain why the smoke bomb went off instead of a horn.

Now back to the fruit. While it is true that the structuring causes merely provided intermediate probabilities for why one fruit ended up at *b* instead of *c*, if a large number of fruits fall from the upstream branches then the structuring causes can do considerable explanatory work. If we find that 1,623 fruits reach *b*, but that only 501 reach *c*, we can explain this pattern in terms of the structuring causes. The structuring causes would predict a *c:d* ratio of approximately 3:1, and the fruit count accords with this prediction, since $1,623:501 \approx 3:1$. Thus, the explanatory and predictive power of structuring causes for cases like this will increase with an increase in the number of events that are triggered (falling fruit, in this example).

We have seen in section 2 that there is a structuring/triggering cause distinction, and in this section that this distinction does not require agents. Structuring causes can play a limited explanatory role for singular events, but can play a much more powerful role for ensembles. It is now time to apply this framework to evolutionary theory in order to see how it sheds light on its causal structure. Let's begin this task by understanding the basic structural building block, the life history.

4. Life History Structures

To understand the relationship between sets of life histories and the structuring/triggering cause distinction, let's begin by considering fitness. Fitness is one of those perennially difficult concepts to define precisely. It has long been recognized that if it is to play a causal-explanatory role in evolutionary theory, fitness should not merely be equated with particular outcomes, like number of offspring produced, or a change in the frequency of a trait. Outcomes do not explain themselves. Instead, if fitness is to play this role, it must be identified with some sort of underlying cause of these outcomes. Brandon [1978] and Mills and Beatty [1979] introduced the propensity interpretation of fitness, in part to solve this problem. Organisms of type *A*, on average, leave more descendants than type *B* individuals because they are fitter, that is, because they bear a propensity to have more offspring.

Although the insights of Brandon, Mills, and Beatty constitute a significant advance in our understanding of evolutionary theory, they leave many lingering questions, such as that of what the ontological basis of fitness is. How, precisely, does fitness do its causal work and what kinds of outcomes can fitness explain? The introduction of my 'block fitness' framework [Ramsey 2006] made some progress on this issue. The block fitness idea is that an organism with a particular genetic constitution and environmental arena can live various possible lives. Its encounter with the environmental heterogeneity can be unlucky, rendering its life short and empty of descendants. Or the organism could live a long, healthy, vigorous life, producing many offspring. There is thus a set of possible life histories, each life history being unique; but the set forms a heterogeneous structure containing patterns of traits and events. And from this set of life histories a single scalar quantity can in theory be derived, which forms the basis of rank-ordered comparisons of fitness values among individuals in a population. (See Pence and Ramsey [2013] for a discussion of the ways in which one can extract a scalar from the set of life histories.) What I hope to accomplish in this paper is to use this framework, in conjunction with the structuring/triggering cause

distinction, to produce a novel account of the causal structure of evolutionary theory.

Consider the set of possible life histories for an individual organism—an orangutan, say, in its forested home in Borneo. If we zoom in on the details of its set of possible life histories, it is easy to see that there is an immensity of ways in which the orangutan can live its life. Many of the small-scale details of the organism's life—looking this way instead of that at a particular moment, scratching itself on an occasion for four seconds instead of five—may have little effect on lifetime outcomes, while other events may have catastrophic effects. The important point, however, is that, although there is an immense number of ways that an orangutan can live its life in the forests of Borneo, there is also an immensity of ways that it cannot live its life. While orangutans may have a diet that includes several hundred food types, they can't—and never will—process their food by, say, cooking or fermenting; they will never hunt and eat other mammals; they will never plant fruits in order to reap a harvest months or years later. These are simply outside of the possible orangutan life histories. This is not to say that these behaviours could not arise in some distant evolutionary future; it is just that, given what orangutans are now, these are not behaviours that exist within their set of possible life histories.

Just as there are both an immense number of ways that an organism can live its life, and at the same time a rather constrained set of possibilities, so it is with the fruit discussed above. A fruit dropped at a (in fig. 2) could, at a particular moment of time, be bobbing up instead of down, be to the left of the centre of the river, or to the right. At this scale of analysis, there is an immense number of possible life histories for the fruit. This does not, however, mean that the fruit could end up anywhere in the universe. Quite the contrary: there are just two places that the fruit can end up—b or c.[3] We have, on the one hand, the constraints of the structuring causes, and, on the other, the immense variety of ways in which the outcomes can be triggered.

The same is the case for the general. There is an immensity of ways that the general might get into his car, lifting the door handle with a particular force, swinging the door open with a specific velocity, inserting the key with a unique flourish. But the fate of the general is nevertheless constrained. In this case, if the general turns the key and is without his family, his life ends. How tightly he grips the key, for example, is irrelevant to the outcome determined by the structuring cause.

It is clear, then, that if one accepts the structuring/triggering cause distinction in the case of the general and the fruit, one should also accept it for biological organisms. For the case of the fruit, it is much clearer what the boundaries are and how the world and properties of the fruit combine to fix the possible life histories for the floating fruit. With organisms, especially animals living in complex environments, the way in which the organism's traits and environment combine to fix the set of possible life histories is staggeringly complex. But it is not different in kind. This life history structure framework, I will now argue, can help us to understand the causal structure of evolution by natural selection, and to show how fitness, selection, and driftability can help causally explain evolutionary outcomes.

---

[3] Of course, if b and c represent lines across the river, not single points, then there are many ways that one can reach b or c (being on one side of the river, moving fast, etc.).

5. Fitness as a Structuring Cause

Consider again the river in figure 2. The structure of the river has a number of features that affect the fruit's movements. The fact that one-quarter of the water goes to the right explains why approximately one-quarter of the fruit dropped ends up at c. But there are other causally efficacious features of the structure as well. For example, the width of the channel affects the water's velocity. If the branch leading to c narrows after it breaks off from the main channel (thereby increasing the water's velocity), then fruit dropped at a would reach c sooner than it would reach b. While there are myriad other features of the structure that can causally affect the fruit outcomes, this one is sufficient in making the general point that we can identify particular features of the structure that can be causally linked to specific features of the outcomes.

Let's now consider organisms and the life history structure described in the previous section. The set of possible life histories for an organism is staggeringly complex and heterogeneous. There is also the fact that the possible life histories are critically dependent on the properties of the organisms. Despite this complexity, however, there are scalar quantities that can be extracted from the set of life histories, and these can play central roles in causal explanations of evolutionary outcomes, just as scalar quantities like channel width can play a central causal explanatory role in the property of travel time for the fruit outcomes.

To connect fitness to the structuring cause idea, consider again the propensity interpretation of fitness. This account as originally formulated [Brandon 1978; Mills and Beatty 1979] was not without problems, since fitness was quantified as the expected offspring production in the next generation. This falls prey to a number of difficulties, like the fact that an increase in the variance in offspring production will lower fitness, even though it does not lower expected number of offspring [Beatty and Finsen 1989]. These problems, however, are not insurmountable. In a recent article [Pence and Ramsey 2013], we offer a new mathematical foundation for the propensity interpretation of fitness, sidestepping these counterexamples. The details of the Pence-Ramsey model need not concern us here, but the model is important in justifying the assumption that each organism's fitness propensity can be quantified as a single scalar, allowing a rank ordering of fitness values for conspecifics in a population.

How should fitness be understood, given the set of life histories described in the previous section? To get a grip on how fitness can be a structuring cause, consider a simple case in which an organism A has only a small number of possible lives. One of these lives ends without reproduction and has a probability of 0.5, and its only other possible life produces a single offspring. Organism A thus effectively flips a fair coin to either reproduce or fail to do so. Compare this to another individual, B, that has a 0.25 probability of having 0 offspring, a 0.5 probability of having 1 offspring, and a probability of 0.25 of having 2 offspring. These possibilities are like the forked river in figure 2: in this case, A has two branches and B has three. In the example of the river, both the characteristics of these paths and the probability of realizing them have important implications for the outcomes of the fruits. So it is with these organisms.

What is the fitness of A and B, and how does fitness causally explain outcomes? Fitness is the probabilistic propensity to survive and reproduce, and it can be quantified via the effects that it has on the structure of the possible life histories. Although it is overly simple (and ultimately problematic, for the reasons mentioned above), let's extract the single scalar for fitness by taking the probability-weighted arithmetic mean

of these values. In this case, the fitness of *A* is 0.5 and that of *B* is 1.0. This scalar quantity can then be used to causally explain evolutionary outcomes (like why a population of 50-50 *A*s and *B*s changed to be only 10 per cent *A*s over a particular time span). This is like being able to extract a single scalar (mean channel width, say) that will predict mean travel time of the fruits from *a* to *c*.

The possible ways in which organisms can live their lives, and the features of the environment that modulate life history outcomes, are much more complex and opaque than the river example. Instead of a simple riverbank, organisms face myriad diseases, dangers, and windfalls as they navigate their lives. The boundaries for organisms are nevertheless just as present as they are for fruit, but they are so much more difficult to see and articulate. This complexity, however, in no way undermines the view presented here. Instead, it just shows that we should not expect simple linear relationships among the variables that affect life outcomes. In section 7, I will elaborate this point and draw the links between fitness and selection. But first let's consider driftability, to see whether it, too, can be a structuring cause of evolution.

6. Driftability as a Structuring Cause

Accounts of drift, like fitness, carry considerable controversy. 'Drift' is often used in ambiguous ways. Sometimes populations are said to change *because of drift*, where drift is singled out as a cause of evolution, and at other times drift merely denotes a kind of outcome, such as a deviation from expectation. The standard biological definition of drift is a change in gene frequency due to random sampling. But it is often left unexplored by biologists what the nature of this random sampling is. Philosophers have taken up this issue, supporting a variety of positions, such as arguing that drift is fundamentally a kind of process [Millstein 2002], a particular sort of cause [Gildenhuys 2009], or a kind of outcome [Brandon 2005].

The account of drift that I will endorse here is one that I have developed elsewhere [2013b]. This account retains the standard definition of genetic drift as used by biologists,[4] but suggests that there is a missing concept—driftability—that can causally explain drift. Within the propensity interpretation framework, fitness is a structuring cause and 'realized fitness' refers to the triggered outcome. Similarly, there is on the one hand *selection*, which is identified as a cause of change, and on the other hand an *evolutionary response to selection*, which is the change precipitated by selection. These are long-standing distinctions [Haldane 1954], but until recently an analogous set of distinctions has not been in play for drift. The driftability/drift distinction parallels that of fitness/realized-fitness. Driftability and fitness are structuring causes; drift and realized fitness are the triggered outcomes.

Consider again the set of possible organismic life histories discussed in section 4. This set has numerous features that can causally explain evolutionary outcomes (just as the river has many properties that can help explain features of the outcome of the fruit's voyage). My claim here is that fitness and driftability operate by structuring the life history possibilities. While fitness is tied to one feature of the constellation of reproductive events along the set of life histories, driftability is tied to another. What, one might ask, are the differences between fitness and driftability, and what effects do they have on population dynamics? The short answer is that, while fitness concerns success and is a

---

[4] Although, because more than genes can drift, I prefer to simply label it 'drift'.

property connected to the probable production of many good quality offspring, driftability concerns the heterogeneity of life history outcomes.

When a population drifts, it evolves in ways that don't accord with fitness rank orderings. The population, as some have put it, exhibits a deviation from expectation [Brandon 2005]. Why did the population drift? The usual answer centres on population size. But population size alone does not fully account for drift. What is required is heterogeneity of the set of life history outcomes for the organisms in the population. If there were no heterogeneity in an organism's set of life histories—if, that is, each organism had only one possible way of living its life—then there could not be any drift (no matter the population size). Like flipping two-headed coins, deviation from expectation (all heads) is impossible independently of how many coins are flipped. Thus, if drift is deviation from expectation, there can't be any drift for organisms with only one possible life history outcome.

Now consider the case in which driftability is not zero, but is low. This could be an example in which there are two possible outcomes but the outcomes do not greatly differ—say, a 0.5 probability of having 7 offspring and a 0.5 probability of having 8. In such a case, the deviation from expectation would be small. The expected value for offspring production in such a case is 7.5, so the largest deviation from this would be 0.5. Note, again, that I have not brought population size into play here. Population size is, however, important for the expression of driftability. In this case, a population of size 1 is guaranteed to deviate from expectation, since no individual can produce 7.5 offspring.[5] In a population of two of these individuals, there is a 50 per cent chance of deviation from expectation. Moving out toward the infinite limit, the probability of there being deviation from expectation approaches zero.

It is easily seen that, as the possible life history outcomes increase, the effect on the deviation from expectation increases. Going from a 0.5 probability of having 7 or 8 offspring to having, say, 5 or 8 offspring will increase the magnitude of the deviation from expectation. This heterogeneity is responsible for populations deviating from expectations, and for this occurring in very significant ways in small populations. And it is driftability that is responsible for this life history heterogeneity. Driftability is a property that structures the set of life histories; it is a structuring cause of evolution.

This interplay of driftability and population size is directly analogous to the interplay between the number of fruit dropped into the river and the expectations of the outcomes. If there is zero heterogeneity in the outcomes of the fruit—if, that is, there are no branches and thus a probability of 1 for the fruit staying in the one and only channel—then there cannot be any deviation from expectation. If three fruits fall into the river, we expect three fruits to end up in the (one and only) channel. The same is true of ten fruits or a hundred. If, on the other hand, the river is branched like that of figure 2, then the outcomes will be a function of both the degree of heterogeneity of the river and the number of fruit dropped. Our expectation for the branch in figure 2 is a 3:1 ratio. If one fruit drops in the river, we would predict a significant deviation from this ratio. With ten fruits, we would predict a closer approximation of the ratio. With 1000 fruits, we would predict a very close approximation. Fruit number, however, is only one component. If the branch heterogeneity were increased, with only a trickle of

---

[5] See Brandon and Carson [1996] for originally making the point that some setups make drift mathematically necessary, not just highly probable.

water passing down the smaller fork, our expectation might be 1000:1 instead of 3:1. It is easily seen that, in such a scenario, many more fruit would have to drop into the river in order to have a good chance of the 1000:1 ratio being approximated. Whereas tens of fruits will do a decent job of approximating the 3:1 ratio, tens of thousands of fruits would be needed for a high probability of the 1000:1 ratio being approximated. Heterogeneity in the structures of life histories thus works in conjunction with the number of fruit to determine probabilities associated with deviations from outcomes. Similarly, driftability—in conjunction with population size—determines the probabilities associated with the deviations from expectation.

7. Selection-Of and Selection-For

Now that we have seen that fitness and driftability are structuring causes of evolutionary outcomes, let's consider how selection fits into this framework. Let's begin by examining one very lucid and influential account of the causal structure of evolutionary theory, that of Sober [1984].

Sober argues that *selection-for* is the central causal concept at the heart of evolution by natural selection. He illustrates his argument with an example of a toy containing a sieve and two kinds of ball—small green balls and larger red ones (the colours have been changed to fit with the fruit example below). The small red balls are small enough to easily pass through the sieve, but the large green balls are too large to easily pass through. Thus, the red balls will be disproportionately represented at the bottom of the toy. Sober's way of describing the situation is this: the small red balls (the objects) were *selected-of*, but it was smallness (the property) that was s*elected-for*. Although the red balls were the ones that ended up easily making it through the sieve, this was not due to the property of being red. It was just that the properties of being red and being small clustered together, allowing the property of being red to tag along.

In contrast to the selection-for concept, Sober argues that fitness is not causal. For him, fitness is like life expectancy, which is a result of putting a lot of unrelated facts together in order to make predictions about how old an individual is going to be when they die. If someone dies at the age of 71, we don't cite their life expectancy as a cause of their death, or as the reason why they died then as opposed to at a different time. The factor that triggered their death may have been one that played a role in the actuarial calculations of their life expectancy, but it was this factor—not their life expectancy—that was responsible for the death. Similarly, or so argues Sober, fitness is a result of myriad causal factors but is not in fact causal.

This appears to contradict the picture of fitness that was described above. Thus, in this section I will offer a framework for understanding how the core concepts of fitness, selection-of, and selection-for are related to one another, and can play a role in causal explanations. To do this, let's return to our fruit example, adding some complexity. Let's say that the overhanging branch drops fruit when the wind whips up, but that the fruit varies in ripeness. The riper fruit is less dense and is red, while the less ripe fruit is denser and green. Above a particular density threshold, the unripe fruit's specific gravity exceeds that of water and hence sinks. The sunken fruit rolls along the riverbed, moving more slowly than the ripe fruit, frequently getting stuck. We might imagine, then, that half of the fruit dropped at *a* in figure 2 is ripe, but that 75 per cent of the fruit that reaches *b* or *c* is ripe.

The river thus exhibits a selection process on the fruit. Using Sober's framework, we would say that the ripe fruit (the objects) are selected-of, but that buoyancy (the property) is selected-for. It happens that the red fruit make it to b and c at a disproportionate rate, but it is not redness that is selected for: buoyancy and redness merely happen to be coextensive traits. Let's now consider fitness and driftability in this case to see if they can causally explain outcomes.

Fitness, again, is a structuring cause of the fruit life histories and concerns successfulness. If making it to point b or c first is what it is to be a successful fruit, then the ripe ones are the more successful. This is not to say that every ripe fruit will be more successful than any less ripe one. It is that a greater proportion of the ripe fruit's possible life histories make a rapid trip down the river. We could imagine, then, an overall success metric for the fruit, which is based on the average travel time of the fruit, an average over its set of possible life histories. This overall structural feature allows us to causally explain the outcomes of fruits of this type. One way to understand why the explanations/predictions become better with larger sets of fruits instead of smaller sets or single fruits is to realize that multiple fruits of the same type (with, therefore, approximately the same set of possible life histories) are exploring more of the space of life histories. Therefore, the fate of a large set of fruits can be better explained or predicted by fitness values since, the larger the set, the higher the probability that the totality of individual fruit life histories is a representative sample of the total set of possible life histories.

While fitness is about success, driftability concerns the heterogeneity of the life histories, in particular the variation in life history outcomes. As pointed out above, as the variation in outcomes increases, we predict larger deviations from expectations. Driftability, then, is tied to intra-organismic variation in life history outcomes. What about selection-of? To understand selection-of, we merely need to move from intra-organismic variation to inter-organismic variation. While driftability concerns intra-organismic variation, selection-of concerns inter-organismic variation in the life histories. Specifically, the focal property is success (that is, fitness): thus, selection-of concerns (intra-generational) inter-organismic variation in fitness. There is selection-of the ripe fruit over the unripe because there is variation in the fitness of the fruit. This is Sober's selection-of: ripe fruit (the objects) are selected-of because there is variation in fitness values, and it is the ripe ones that are fitter.

Selection-of in this sense is thus not a process; it is constituted by fitness differences. Evolution by natural selection, however, is a process. And it is one whose outcomes we can explain via fitness differences (that is, selection-of). The life history framework thus unifies fitness, driftability, and selection-of. In this discussion of selection, the focus has been on organisms (objects), not on their traits (properties). Let's now consider properties and how to understand the selection-for concept.

If there are fitness differences among individuals (and therefore selection), then there must be differences in their set of possible life histories due to the traits that the organisms will bear if such a life history is realized. Some of these life history traits will be the ones that are responsible for the fitness differences. Among the traits responsible for the fitness differences, the ones that make a positive difference for fitness are the ones that are selected-for. Let's consider how this understanding of the selection-for concept would play out for our fruit example before considering more of its implications for organisms.

Recall that fruits dropped at *a* vary in their colour and density, the riper ones being red and buoyant enough to float. The red fruit (the objects) are thus selected-of, but what is selected-for? The answer to this question concerns which life history properties make the difference between making it down the river quickly or not making it down quickly (or at all). The difference in this case is due to the density of the fruit. It is thus low density (or high buoyancy) that is selected-for.

It is important to distinguish between traits that are coextensive with those making fitness differences, and those responsible for those differences. Fruit colour correlates with speed of movement downstream as well as the probability of making it all the way down. It is not, however, a trait selected-for: the trait is not responsible for changing the set of life histories in fitness-affecting ways; it is merely riding along. How could we know this? We might proceed as follows. We could make observations about how the fruits behave as they make their way down the river. We could note that the difference in outcome seems to be tied to whether or not the fruit floats, and note that it is the density of the fruits that determines whether or not they float. This conjecture could then be tested. For example, fruit could be painted and it could be seen whether colour changes affect their outcomes. Or plastic artificial fruit could be created that vary in their size, density, and colour—and these could be dropped at *a* to see which properties affect their outcomes in which ways.

It is important to emphasize that selection-for, as described here, is not about outcomes: it is not a way of categorizing the results of evolution. Instead, it is about properties that change the possibility space in fitness-affecting ways. If, for example, there is an unlucky ripe fruit that happens to spend a lot of time near the slow-running edge of the river, it may lose the race to an unripe fruit that happened to get a luckily obstruction-free path down the middle of the riverbed. It is a mistake to examine the fates of these two fruits and say that the denser one was in this case selected-for. This is wrong because it is basing the answer to the question of whether there was selection-for on a single life history, not on the totality of life histories. Instead, for a property to be selected-for, it need only make a positive difference to fitness. It must only, that is, be responsible for a difference in the overall structure of possible life histories in a positive fitness-affecting way—it need not make a positive difference for each life history.

The selection-for concept works for organisms just as it does for the fruit in the river. Consider the arctic hare, *Lepus arcticus*. In the far north, the hare is white all year around. But in the southern part of its range, where the summers are longer, it changes colour in the summer to better match its surroundings. If we find that some of the southern hares retain their white pelts through the summer, and that they have a higher predation rate, we can conclude that the individuals that change colour are selected-of. If we confirm that it is the lack of camouflage that underlies the fitness differences, then it is the trait of moulting into a darker colour during the summer that is selected-for. Having the trait of moulting into the darker pelt helps to shape the set of possible life histories for the hares. The white ones have a greater proportion of their life histories end in predation. And it is this structural difference in the life history space that underlies the difference in outcomes for the hares.

8. Conclusions

The theory of evolution by natural selection provides a basis for how stability and directional change can be caused and, perhaps most centrally, provides a naturalistic

account of the origin and maintenance of adaptations—traits with apparent design and purpose. This much is clear. What has remained unclear is how, precisely, this theory works and which parts of it are causal. This has led some to argue that its central concepts should not in fact be given a causal construal [Matthen and Ariew 2002; Walsh, Lewens, and Ariew 2002; Pigliucci and Kaplan 2006], or that only a single concept—selection-for, for example—is causal [Sober 1984]. This paper has attempted to provide a general framework with which to understand the causal nature of evolution by natural selection. Fitness and driftability are causes, although they are causes of a special sort: structuring causes. It is my hope that this framework will provide a useful platform for the causalists or, perhaps, a clearer target for the statisticalists. Playing either role, it should help to advance the debate and the shared project of having a deeper understanding of evolutionary theory.[6]

ORCID

Grant Ramsey http://orcid.org/0000-0002-8712-5521

References

Ariew, André and Zachary Ernst 2009. What Fitness Can't Be, *Erkenntnis* 71/3: 289−301.

Beatty, John and Susan Finsen 1989. Rethinking the Propensity Interpretation of Fitness: A Peek Inside Pandora's Box, in *What the Philosophy of Biology Is: Essays Dedicated to David Hull*, ed. Michael Ruse, Dordrecht: Kluwer Academic Publishers: 17−30.

Bouchard, Frédéric and Alex Rosenberg 2004. Fitness, Probability, and the Principles of Natural Selection, *The British Journal for the Philosophy of Science* 55/4: 693−712.

Brandon, Robert N. 1990. *Adaptation and Environment*, Princeton: Princeton University Press.

Brandon, R.N. 1978. Adaptation and Evolutionary Theory, *Studies in History and Philosophy of Science Part A* 9/3: 181−206.

Brandon, Robert N. 2005. The Difference Between Selection and Drift: A Reply to Millstein, *Biology and Philosophy* 20/1: 153−170.

Brandon, Robert N. and Scott Carson 1996. The Indeterministic Character of Evolutionary Theory: No 'No Hidden Variables Proof' but No Room for Determinism Either, *Philosophy of Science* 63/3: 315−37.

Brandon, Robert N. and Grant Ramsey 2007. What's Wrong with the Emergentist Statistical Interpretation of Natural Selection and Random Drift?, in *The Cambridge Companion to the Philosophy of Biology*, ed. by David L. Hull and Michael Ruse, Cambridge: Cambridge University Press: 66−84.

Dretske, Fred 1988. *Explaining Behavior: Reasons in a World of Causes*, Cambridge, MA: The MIT Press.

Dretske, Fred 2004. Psychological vs. Biological Explanations of Behavior, *Behavior and Philosophy* 32/1: 167−77.

Gildenhuys, Peter 2009. An Explication of the Causal Dimension of Drift, *The British Journal for the Philosophy of Science* 60/3: 521−55.

Haldane, J.B.S. 1954. The Measurement of Natural Selection, *Caryologia* 6. Suppl.: 480−7.

Jackson, Frank and Philip Pettit 1992. In Defense of Explanatory Ecumenicalism, *Economics and Philosophy* 8/1: 1−21.

Mackie, J.L. 1965. Causes and Conditions. *American Philosophical Quarterly* 2/4: 245−64.

Matthen, Mohan and André Ariew 2002. Two Ways of Thinking about Fitness and Natural Selection, *The Journal of Philosophy* 99/2: 55−83.

Mills, Susan K. and John H. Beatty 1979. The Propensity Interpretation of Fitness, *Philosophy of Science* 46/2: 263−86.

Millstein, Roberta L. 2002. Are Random Drift and Natural Selection Conceptually Distinct? *Biology and Philosophy* 17/1: 33−53.

Millstein, Roberta L. 2006. Natural Selection as a Population-Level Causal Process, *The British Journal for the Philosophy of Science* 57/4: 627−53.

Pence, Charles H. and Grant Ramsey 2013. A New Foundation for the Propensity Interpretation of Fitness, *The British Journal for the Philosophy of Science* 64/4: 851−81.

Pigliucci, Massimo and Jonathan Kaplan 2006. *Making Sense of Evolution: The Conceptual Foundations of Evolutionary Biology*, Chicago: University of Chicago Press.

Ramsey, Grant 2006. Block Fitness, *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 37/3: 484−98.

Ramsey, Grant 2013a. Can Fitness Differences Be a Cause of Evolution? *Philosophy & Theory in Biology* 5: e401.

Ramsey, Grant 2013b. Driftability, *Synthese* 190/17: 3909−28.

Ramsey, Grant 2013c. Organisms, Traits, and Population Subdivisions: Two Arguments Against the Causal Conception of Fitness? *The British Journal for the Philosophy of Science* 64/3: 589−608.

Sober, Elliott 1984. *The Nature of Selection*, Chicago: University of Chicago Press.

Sterelny, Kim 1996. Explanatory Pluralism in Evolutionary Biology, *Biology and Philosophy* 11/2: 193−214.

Walsh, Denis M. 2007. The Pomp of Superfluous Causes: The Interpretation of Evolutionary Theory, *Philosophy of Science* 74/3: 281−303.

Walsh, Denis M. 2010. Not a Sure Thing: Fitness, Probability, and Causation, *Philosophy of Science* 77/2: 147−71.

Walsh, Denis M., Tim Lewens, and André Ariew 2002. The Trials of Life: Natural Selection and Random Drift, *Philosophy of Science* 69/3: 429−46.