Biol Theory
DOI 10.1007/s13752-016-0254-2

ORIGINAL ARTICLE

Explaining Drift from a Deterministic Setting

Pierrick Bourrat[1,2]

Received: 25 September 2016 / Accepted: 21 November 2016

Abstract Drift is often characterized in statistical terms. Yet such a purely statistical characterization is ambiguous for it can accept multiple physical interpretations. Because of this ambiguity it is important to distinguish what sorts of processes can lead to this statistical phenomenon. After presenting a physical interpretation of drift originating from the most popular interpretation of fitness, namely the propensity interpretation, I propose a different one starting from an analysis of the concept of drift made by Godfrey-Smith. Further on, I show how my interpretation relates to previous attempts to make sense of the notion of expected value in deterministic setups. The upshot of my analysis is a physical conception of drift that is compatible with both a deterministic and indeterministic world.

Keywords Determinism · Drift · Evolution · Fitness · Probabilities

Introduction

Drift is often regarded as one of the forces or causes for evolutionary change, where evolutionary change in turn is regarded as a change in frequency of the different variants composing a population (Sober 1984, p. 34; Gillespie 2004; Stephens 2004). This force or cause is classically distinguished from natural selection (Sober 1984; Millstein 2002) and has been defined as the process(es) by which evolutionary change occurs by "chance" or "accident," whereas natural selection is a process by which evolutionary change occurs because some variants have an advantage over others. Although there is no consensus on what drift is (Millstein 2016), with the meaning of drift today having been extended from its historical definition that developed in the context of population genetics, there does seem to be an agreement that chance is an important feature of drift (Okasha 2006, p. 32; Millstein 2016). Following Okasha (2006, pp. 32–33), a useful starting point to characterize chance and accident in statistical terms is to consider them as being deviations from expected reproductive outputs or expected fitness.[1] Once characterized as such, evolutionary change due to natural selection results from differences in expected reproductive outputs, and evolutionary change due to drift results from the random deviations from expected reproductive outputs.

Although the notion of expected reproductive output can be used quite straightforwardly in statistical models and general equations of evolutionary change such as the Price equation, how one should interpret physically these mathematical entities in the context of evolutionary theory is quite challenging. The main motivation for a physical interpretation of drift is that a purely statistical notion of drift will be unsatisfactory if one wants to distinguish two or more distinct types of causal processes that can lead to the very same observed phenomenon. Obviously, assessing

✉ Pierrick Bourrat
p.bourrat@gmail.com

1 Department of Philosophy, History and Philosophy of Science Unit & Charles Perkins Centre, University of Sydney, Sydney, NSW, Australia

2 Department of Philosophy, Macquarie University, Sydney, NSW, Australia

---

[1] Note that this way of characterizing drift might not capture all the different meanings of drift used in the biological literature, but it captures some important meanings of it and is grounded in influential theoretical work (e.g., Grafen 2000; Rice 2004, 2008). For thorough reviews of the notion of drift see Plutynski (2007) and Millstein (2016).

the concept of drift from a purely phenomenological/statistical perspective will not allow one to make such a separation. It seems therefore desirable to have a concept of drift in which deviations from expected values are explained physically rather than assumed as mathematical/statistical truths. Note that I do not imply here that purely statistical notions of drift should be abandoned. Rather I claim that it is reasonable to look for a physical conception of drift.

What is known as the propensity interpretation of fitness (Brandon 1978; Mills and Beatty 1979; Beatty and Finsen 1989; Sober 2001) permits one physical interpretation of drift. An expected reproductive output or fitness, under this interpretation, is viewed as a tendency or disposition comparable to familiar examples like fragility.[2] Entities[3] have a disposition to produce a certain number of offspring in the same way glass has a disposition to break. Evolutionary change due to drift increases proportionally with the deviations from expected reproductive outputs resulting from the dispositional properties of entities to produce offspring.

However, propensity interpretations of probability in general or when applied to fitness are controversial. In fact, propensity interpretations have been claimed to be empty accounts of probability because it is not clear what propensities represent (Eagle 2004; Hájek 2012). Furthermore, it is classically assumed that the least problematic propensity interpretation of probabilities, namely single-case propensities, requires indeterminism (Millstein 2003; Rosenthal 2010). If fitnesses are (single-case) propensities and they play a causal role in evolution, they are simply brute indeterministic facts about the physical reality (Bouchard and Rosenberg 2004). Yet, it is conceivable that in many cases (maybe the majority of them) what is referred to as drift is occurring in a fully deterministic setup, or sufficiently close to one. At any rate, an increasing number of philosophers believe that indeterminism is eliminable from evolutionary theory (see Weber 2001; Sober 2010 for discussions), which questions the adequacy of propensity interpretations of fitness and the physical interpretation of drift that derives from them.[4] It is perfectly reasonable to have a concept of fitness that relies on propensities and ultimately indeterminism, but it is more controversial to apply it when there is no reason to suppose indeterminism and strong reasons to suppose determinism (although for a defense of the indeterminism thesis see Brandon and Carson 1996; Glymour 2001; Stamos 2001). In such cases the distribution of reproductive outputs of entities of the same type and the deviations from expected values must necessarily be accounted for by invoking other factors than propensities. Small differences in initial conditions of the environment seem, prima facie, to be a good candidate, and it is the type of factor I will explore and relate in this article.

In light of the above remarks it seems reasonable to demand a physical interpretation of drift satisfying two related desiderata. First, it is desirable to have a physical account of drift that does not rely exclusively on one kind of interpretation of fitness, namely the propensity interpretation, especially because it is grounded, as we have seen, in one of the most controversial interpretations of probability. Second, and related to the previous point, our account should be compatible both with a deterministic and an indeterministic world. This is because even if in some cases indeterministic processes seem to have consequences on reproductive output (see, for instance, Glymour 2001), it is plausible that many cases of drift are straightforward cases of determinism, or at least that they can be considered as such. This resonates with Millstein's (2003) remark that probability in evolutionary theory should be compatible with both determinism and indeterminism. I will assume throughout a deterministic setup of which the initial conditions can be considered as either deterministic or indeterministic. Thus my account will be compatible with both a deterministic and an indeterministic world.[5] The main aim of this article is to develop an alternative or at least complementary physical interpretation of drift to the interpretation borne out of the propensity interpretation of fitness and that satisfies the two desiderata.

The article will run as follows. In the next section, I present Godfrey-Smith's (2009) framework for drift, which happens to be compatible with the two desiderata. More

---

[2] Note that strictly speaking expected reproductive outputs are just one way to characterize propensities. In simple models propensities can be reasonably approximated with expected value (e.g., Mills and Beatty 1979). But doing so leads to a number of problems (see, for instance, Beatty and Finsen 1989). A number of authors have proposed new propensity interpretations of fitness (e.g., Brandon 1990; Ramsey 2006; Pence and Ramsey 2013) or some other interpretation (e.g., Abrams 2009) that attempt to solve these problems.

[3] By "entity" I will mean any object of a population able to undergo evolution by natural selection (e.g., organism, gene, cell, group).

[4] Note that some could consider that small differences in the environment affecting reproductive outputs are in some sense the

Footnote 4 continued
"brute indeterministic facts about reality" even though the setup is deterministic. This will not be the interpretation given here since by definition indeterministic facts cannot occur in a deterministic setup. I recognize that epistemically one might want to take some facts as indeterministic even though the setup is deterministic, but the epistemic and the ontological questions should not be confused.

[5] I am appealing here to the same kind of distinction made by Rosenberg (2001, pp. 537–538) between a statistical or indeterministic theory and an indeterministic world. Having a theory that is indeterministic does not imply that the world is and vice versa. Here the relevant distinction is that although a system might be deterministic, its input might not, and vice versa.

particularly I present his view that drift results on the one hand from differences in reproductive output due to differences in extrinsic properties as opposed to difference in intrinsic properties (which should be attributed to natural selection), and on the other hand from the population exhibiting what Godfrey-Smith calls a "low continuity." In the section following, I demonstrate that although it is on the right track, this framework is problematic for a number of reasons. Starting from Godfrey-Smith's framework, I present my physical account of drift. This account is fundamentally independent from probabilities but makes it possible to explain them in physical terms in classical models of evolutionary change. In the last section, I respond to some objections one might have with my framework and show that under some particular conditions I briefly describe, it is compatible with a new objective interpretation of probability in deterministic setups. I will call this interpretation "natural-range interpretation of probability" following Rosenthal (2010).

Godfrey-Smith on Drift

In his recent book, Godfrey-Smith (2009, pp. 53–63) develops a principled way of distinguishing natural selection from drift, from the perspective of entities forming a population. His framework happens to be compatible, at least in part, with the view that the distribution of reproductive outputs of the entities of a same type is due to differences in initial conditions. Most importantly it is the only physical account of drift I am aware of–besides the more recent ones of Ramsey (2013) and Strevens (2016), the latter of which takes Godfrey-Smith's account as a source of inspiration—that satisfies the two desiderata listed in the introduction. Namely, it does not rely on the propensity interpretation of fitness, and it is compatible both with determinism and indeterminism. For that reason it is a natural place for me to start. Godfrey-Smith holds the view that the distinction between natural selection and drift has something to do, among other things (more on this below), with the notions of intrinsic and extrinsic properties of the entities that make up populations.

Godfrey-Smith proposes that when a population contains variation in intrinsic properties between its members, and this leads them to have different reproductive outputs, the resulting evolutionary change should be attributed to natural selection. Conversely, when differences in reproductive output are due to differences in extrinsic properties, the evolutionary change resulting should be attributed to drift. He defines an intrinsic property as a property that, in contrast to an extrinsic one, does not depend on the existence and arrangement of other objects. A good example of intrinsic property for an organism is having mitochondria. Examples of extrinsic properties include being at a particular location or someone's sibling.[6]

The rationale behind this view is that when intrinsic properties, as opposed to extrinsic ones, are causally responsible for differences in reproductive outputs, they can systematically be attributed to their bearers. In some sense, intrinsic properties are constitutive of an entity while extrinsic properties are not. Another way to understand this distinction is to use counterfactual dependences. When evolutionary change is due to drift, according to Godfrey-Smith, then had some of the circumstances of the entities in the population been different, the extrinsic properties of those entities would have differed in a way that led to different reproductive outputs. When evolutionary change is due to natural selection, then had the entities themselves been otherwise (because of differences in intrinsic properties), their reproductive outputs would have been different, irrespective of the circumstances of those entities. By associating extrinsicness with drift, we recover the notion of accident classically associated with drift, where accident can be understood as "that which does not causally depend on the entities forming the population." It follows that two types of entities with different extrinsic properties (that do not ultimately causally depend on intrinsic properties), even in a fully deterministic setup, should exhibit drift. Under this framework, the probabilistic nature of a type's reproductive output should not necessarily be regarded as the result of indeterminism, but as a measure of the extent to which the entities of a type differ in extrinsic properties that lead to differences in reproductive outputs. These differences might be deterministic in nature or not.

The fact that differences in reproductive outputs stem from extrinsic properties does not necessarily imply that reproductive outputs are probabilistic in nature. In fact, it follows from Godfrey-Smith's interpretation that had each extrinsic property causing a difference in reproductive output between types been the same for all the members of a type (if such a thing was possible), each member of a given type (supposing these have exactly same intrinsic properties) would have had the same reproductive output.[7] Furthermore, unlike the classical probabilistic account of

---

[6] Although, as Godfrey-Smith recognizes, the terms "intrinsic" and "extrinsic" are controversial in philosophy, I will follow him in his view that the main idea behind the distinction can be useful. For more on this distinction see Weatherson (2014).

[7] This consequence relies on the assumption that no asymmetric extrinsic properties (e.g., "being less tall than") are involved in differences in reproductive outputs, in which case it would be impossible to have the same extrinsic property for all the members of a population. This assumption is of course far from any real biological case, but I use it merely to make a conceptual distinction.

drift, this view of drift does not *imply* that with a perfectly identical selective environment[8] for each entity of a population, smaller populations will exhibit higher levels of drift. However, using this view, one can *explain* why reproductive outputs of types appear probabilistic and why drift decreases as population size increases even in a deterministic setting.

To see this, let us suppose a population composed of two types of entities that do not differ in intrinsic properties in a way that has consequences on reproductive output. Each can be found with a given probability in several different types of (micro)states of the environment.[9] For a given entity, each type of microstate is supposed to lead to different (but the same within each type) reproductive-output consequences (which might be highly variable). If a population is composed of a small number of entities and there is nothing intrinsic to the types that determines in which microstate they find themselves,[10] it is very unlikely that for each entity of a given type in a particular type of microstate, there will be another entity of the other type in the same type of microstate (of course this probability will depend on the number of microstates in the environment).[11] However, the probability of finding matching entities of each type in the same type of environmental microstate will increase as population size increases (also subject to the number of microstates), assuming there is no reason why one entity would be found in one microstate rather than another one. When the population size is infinite, any given entity of one type in a particular environmental microstate has a matching entity of the other type in the same environmental microstate (assuming a finite number of microstates). At that point, no significant differences in reproductive outputs between the two types should be observed, and thus no drift should occur.

But there are cases in which a significant level of drift can occur even when the population is large. This often happens, for instance, when the environmental variation leading to differences in reproductive outputs is *not* distributed evenly in the habitat. In population genetics this phenomenon is captured by the notion of effective population size. Drift, population geneticists tell us, occurs when the *effective* rather than the *actual* population size is small. The effective population size corresponds roughly to the number of individuals that *effectively* contribute to the next generation, but it is assessed by comparing the real population to an idealized population meeting the assumptions of the Wright-Fisher model in which every individual has the same chance to contribute offspring to the next generation (Hamilton 2009, p. 73). Does my interpretation of Godfrey-Smith's account allow the notion of effective population size to be recovered? Better than that, it gives a physical interpretation of it.

Suppose a case, similar to the previous one, in which only one or several microstates, say a few patches, are "viable" while all the others are not. A biological example matching this case would be dandelion seeds (produced in a large numbers) blown by the wind in an environment that is mostly bare rock.[12] In this case, in spite of the large population sizes of seeds of two types that do not differ in intrinsic properties affecting reproductive output, evolutionary change could easily lead one type (A) or the other (B) to be fixed in the population in a "drifty" way. The reason why the evolutionary change is drifty here is because there is a large ecological variation between each microstate of the environment, not because the population size is small. To see this, suppose our two types of dandelion seeds to be in equally large numbers, and there are only two favorable patches accessible in the population, each of which allows for only one seed to grow; then in 50% of the cases seeds of the same type will colonize both patches, which would be associated with drift (as opposed to a case in which each type has access to one favorable patch), since at the next generation one type will be fixed in the population. When this happens, the reason why one type becomes fixed in the population is because it happened to be found more often than the other type in a favorable type of microstate.

Cases like the dandelion seeds demonstrate why environmental heterogeneity is also an important factor with respect to drift (Bouchard and Rosenberg 2004; Pfeifer 2005; Ramsey 2013). And this is one of the reasons why effective rather than actual population size is used to measure drift in population genetics. In fact, as I mentioned earlier, population genetics does not consider the environment as a parameter influencing reproductive outputs, so it uses the notion of effective population size to get around the problem by assuming that the actual population studied is an idealized population in which all the entities *would* actually have had the same opportunity to reproduce given to them. Thus, with the notion of

---

[8] I use here Brandon's (1990) notion of selective environment as all the factors surrounding an organism (or more generally an entity) that differentially affect its reproductive output when compared to another organism.

[9] This way of representing the environment in the context of evolutionary change is compatible with Abrams's (2014) framework on environmental variation.

[10] Suppose, for instance, a case with ten possible microstates present in equal proportions. It would imply that entities encounter each microstate with equal probability (0.1).

[11] This explanation is to put in perspective the account proposed by Ramsey (2013), for whom drift results from the heterogeneity in the possible causes one entity can experience, leading to different reproductive outputs.

[12] Thanks to Kim Sterelny for bringing this case to my attention.

effective population size, ecological differences are "emulated" by population size differences (Hamilton 2009, p. 73). Thus my interpretation of Godfrey-Smith's framework in terms of differences in extrinsic properties predicts that drift will be negligible as long as each type of environmental (micro)state is equally accessed by each type of entity in a population. If not, then differences in extrinsic properties will be responsible for differences in reproductive output, which will produce evolutionary changes in a drifty way.

Differences in access of types of microstate by type of entities can be obtained in three different ways. First, they can arise directly from a small population size when compared to the number of environmental microstates (there are not enough entities to "fill" all the possible states in the environment). Second, they can arise when the number of one type is much smaller than the others, in spite of the population size being large. This can explain why, for instance, new beneficial mutations are largely driven by drift (Gillespie 2004, p. 21). Finally, they can come from the number of instances of some microstates being too small to allow each type of entity to have equal access to that state. This corresponds, for instance, to the dandelion case presented above.

We can see here that one difference between the classical statistical view and this new physical or causal view of drift is that the latter, contrary to the former, does not take probability distributions as given, but rather explains how one can obtain them when considering the differences in extrinsic properties of the entities that matter for reproduction. Toward the end of the "Responses to Objections" Section, I provide details on how this interpretation of drift can be linked to probability, using—following Rosenthal (2010)—what I call the natural-range interpretation of probability.

Thus, following my interpretation of Godfrey-Smith's framework, if the environment is patchy (and even more so if it is patchy in an unpredictable way), we can explain why not all entities of a type have the same reproductive output: simply because differences in extrinsic properties lead to differences in reproductive outputs between the members of the types. When these differences are on average different between two types, we obtain drift.

Godfrey-Smith's (2009) view on natural selection and drift is actually more complex than the one I have presented so far. The dependence of reproductive output on intrinsic properties is for Godfrey-Smith only one of at least five important parameters to characterize Darwinian populations, that is, populations able to exhibit evolution by natural selection.[13] Godfrey-Smith uses the symbol S to denote the degree of dependence of realized fitness[14] on intrinsic properties. The higher S the more differences in reproductive output between the members of a population depend on differences in intrinsic properties and should be associated with natural selection. Along with S the four other parameters are fidelity of heredity H, abundance of variation V, competitive interaction with respect to reproduction α and continuity C (for more on H, V, and α see Godfrey-Smith 2009, pp. 44–53).

For Godfrey-Smith the parameter C also explains drift. It represents the level of change in reproductive output induced by small changes in an entity's phenotype (2009, p. 57), which can be viewed in some cases as the degree of smoothness of the fitness landscape. A population scoring high on C means, in some cases, that the fitness landscape is very smooth. For more on this parameter see Godfrey-Smith (2009, pp. 57–59) and the next section. Godfrey-Smith claims that a population having a low C, so that small changes in the phenotype lead to large changes in reproductive output, should be associated with drift. Thus, for him, the paradigmatic cases of drift result from a combination of low C and low S, but low C without low S and low S without low C can be independently associated with drift. He writes: "What looks *most* like drift is low C and low S. But cases where only C is low [without low S]—cases where tiny internal accidents lead to reproductive consequences—might also look like drift. And there is a yet more attenuated sense in which low S without low C can look like drift" (Godfrey-Smith 2009, p. 61). It should be noted that an entity's phenotype will often depend on both its extrinsic and intrinsic character, so there seems to be from the outset a contradiction in Godfrey-Smith's view in which drift could at the same time depend on differences in extrinsic and intrinsic character. I will come back to this apparent contradiction in the next section.

Beyond Godfrey-Smith's Distinctions

Although I find Godfrey-Smith's distinctions useful, I will argue here that his treatment of drift is not entirely successful. My first point of disagreement is that Godfrey-Smith considers that small differences in "everything about an organism [read 'entity']" (2009, p. 61), that is, both intrinsic and extrinsic properties, can in some cases lead to large differences in reproductive outputs (and hence low C), and should thus be associated with drift. I find this problematic for two reasons. First, I am inclined to consider that *large* differences in reproductive outputs due to *small* differences in extrinsic properties merely represent a

---

[13] According to Godfrey-Smith (2009, p. 63), this list of five parameters is incomplete and could include other features.

[14] For our purposes in this article this is equivalent to realized reproductive output.

subset of the cases in which differences in extrinsic properties lead to differences in reproductive outputs. Different levels of C, when they concern differences in extrinsic properties, I claim, *modulate* or change the level of difference in reproductive output resulting from differences in extrinsic properties, but they do not qualitatively change the nature of the processes occurring in the population. And in fact, I explained the case of the dandelion in the previous section merely by making references to differences in extrinsic properties when these differences (which could be very small) are responsible for large differences in reproductive output.

Second, I find even more problematic Godfrey-Smith's proposition that small differences in intrinsic properties leading to large differences in reproductive outputs could be associated with drift. In fact, these differences should be regarded, following his own distinction with respect to the parameter S, as differences that should be attributed more naturally to natural selection than to drift. If Godfrey-Smith's distinction between intrinsic and extrinsic properties is right, it cannot be the case that small differences in intrinsic properties are sometimes attributed to natural selection and sometimes attributed to drift simply because these differences lead to greater differences in reproductive outputs, especially if one is interested in giving an interpretation of drift from a causal or physical rather than statistical perspective, which is my aim and which seems to be at least partly Godfrey-Smith's aim.

Although I think the parameter C in and of itself should not be associated with drift, or only inasmuch as it amplifies the effects brought about by differences in extrinsic properties (that is, associated with S), I regard C as a useful concept for understanding Darwinian populations. I propose that C when it concerns small differences in intrinsic properties (C_i as opposed to C_e that concerns small differences in extrinsic properties) helps to conceptualize the conditions under which natural selection can lead to complex adaptations. If slight differences in intrinsic properties lead to dramatic changes in reproductive outputs, that is, if the population exhibits a low C_i, it is very unlikely that anything like a lensed eye could ever evolve by natural selection, even though natural selection is present in the population. If one uses the fitness landscape metaphor (although see Godfrey-Smith 2009, pp. 57–59 for the limitations of this metaphor), a low C involves a very rugged landscape in which a population cannot smoothly ascend to a peak. As a result there is selection but it is not directional, and no pattern (such as one of diversification or stabilization) can be detected. Thus different entities, although phenotypically very similar, end up at very different places on the landscape.

Thus in my view, Godfrey-Smith's project, especially with his Darwinian space (2009, pp. 63–67), is a (successful) attempt to delimit the conditions under which natural selection will lead to adaptations as classically understood by evolutionary biologists, that is, through cumulative evolution. Yet, the question of whether natural selection can lead to complex adaptations in a population and the question of whether natural selection, as opposed to drift, is occurring in that population are quite different and should be distinguished. The process of natural selection does not necessarily lead to the production of complex adaptations or even to evolutionary change, as has been emphasized by Godfrey-Smith himself (Godfrey-Smith 2007, 2009, pp. 24–26). It is also standard in quantitative genetics when using the Breeder's equation, to separate the *response to selection*, that is, evolution by natural selection, from *selection* (see Falconer 1981).[15] I suggest that cases of populations with low C_i look similar to cases with a low response to selection in the sense that natural selection is occurring but it does not lead to complex adaptations, because heritability is low. Even if evolution by natural selection without complex adaptations might be extremely hard to distinguish empirically from evolution due to drift—that is, following my framework, distinguishing a low S and a low continuity concerning extrinsic factors (low C_e), from a high S and a low C_i—the conceptual distinction is important.

There is also a more general problem with Godfrey-Smith's framework and more particularly the intrinsic/extrinsic distinction, which is that it does not cover a number of cases that ought to be counted as drift. This is not a fatal problem for the framework, but it renders it incomplete. One way to appreciate this incompleteness is by remarking that any biological property, say for instance "height," is diachronically the result of the interaction between the bearer of the property and its environment. Had a given organism been put in a different environment from birth, its height might have been very different. Godfrey-Smith's distinction between intrinsic and extrinsic properties only accounts for "synchronic" dependences on reproductive output, leaving out "diachronic" ones. Yet as I show below, diachronic dependences on reproductive output matter a lot with respect to natural selection and drift (see also Bourrat 2015a).

Consider the following intrinsic property of an organism, "amount of fat." The amount of fat contained by each organism of a population of, say, any mammals, is generally different, and it is reasonable to suppose that this might have consequences on their reproductive output. Using Godfrey-Smith's framework, all the differences in reproductive output due to differences in amount of fat

---

[15] Specifically, the Breeder's equation tells us that the response to selection R is equal to a selection differential S multiplied by heritability h^2 so that $R = S \times h^2$.

contained by organisms should be attributed to natural selection because "amount of fat" is an intrinsic property. The problem here is that there are cases in which containing a certain amount of fat leading to a different reproductive output, when compared to another organism, should intuitively be attributed to drift. Imagine, for example, that two organisms have different reproductive outputs due to the fact they contain a different amount of fat, but the difference here is the result of different life histories that cannot causally be traced back to any of their intrinsic properties. For instance, suppose that the two organisms have the same susceptibility to a disease *D*. Yet, for some reason one gets *D* and has to spend more energy to eliminate it. To do so it burns a larger amount of fat than the other organism. As a result, the two organisms have different amounts of fat and produce different numbers of offspring.

This sort of case should be considered as a case of drift because the phenomenon seems to be the result of a chance event (one organism got the disease while the other did not). Yet, Godfrey-Smith's distinction is blind to them. Another more dramatic way to make the same point is to take the case in which one organism is burnt (which will usually[16] have bad consequences for its reproductive output) while the other is not. Because there is a difference in intrinsic properties between the two organisms leading to a difference in reproductive output, using Godfrey-Smith's distinction would lead to the view that the difference in reproductive output should be attributed to natural selection. But suppose two scenarios: in one the difference between the burnt and non-burnt organism is due to the fact they have two different alleles, one of which predisposes the organism to escape from bushfires, while the other does not. In the other scenario, the difference is due to one organism being struck by lightning while the other happily survives and reproduces.[17] In the first scenario there is no question that the difference should be attributed to natural selection, but in the second differences in both extrinsic properties (location) and intrinsic ones (whether the organism is burnt or not) are associated with a difference in reproductive output. In the second scenario, however, the difference in intrinsic properties is fully explained by differences in extrinsic properties. Godfrey-Smith's framework does not straightforwardly imply that this is a case of drift. This is problematic since this example is considered as a paradigmatic case of drift. It demonstrates that Godfrey-Smith's distinction does not account for all cases of drift and can be unclear even in such paradigmatic cases.

But all is not lost. I agree with Godfrey-Smith that to be associated with natural selection, a difference in properties between members of a population must be intrinsic (or causally determined by a difference in intrinsic properties). But this is insufficient since the simple and biologically plausible examples presented above are clear counterexamples. To complete Godfrey-Smith's framework, I propose thus that properties must also be *invariable* or insensitive to differences in extrinsic properties (for a response to the different worries one might have with the notion of invariability see the next section). The differences between being fat or not and being burnt (due to the lightning) or not in the examples above, although they are due to intrinsic properties, can vary over time or depending on the circumstances, because they are ultimately due to differences in the environment the organisms were subjected to (differences in extrinsic properties). Yet, because the initial environmental conditions an organism is subjected to do not depend in these cases on its biology (they are accidental), these cases should be associated with drift.

If the reasoning above is correct, intrinsic properties should therefore be decomposed into two sub-properties that will allow us to account for the more complex cases in which intrinsic properties vary over time: namely, intrinsic-*invariable* properties, such as having a particular gene or having two legs; and intrinsic-*variable* properties, such as an amount of fat or being burnt due to a particular life history causally independent from any intrinsic-invariable properties of the entity. Both intrinsic-variable and intrinsic-invariable properties should be understood while specifying a range of environmental conditions and over a given period of time, that is, over a range of possible extrinsic properties. This is crucial since what is invariable here and now might not be at a later time or under different conditions. This means that natural selection and drift are context sensitive. It is possible to imagine that a property such as height, for instance, that does not vary under a range of specific conditions and can thus be considered as invariable in those conditions, would do so under another set of conditions. For instance, if the height of some organisms of a population was measured after they were subjected to a different degree of gravitational force or over a different period of time, the height would have changed, and this might have consequences on their reproductive output. Everything else being equal, this difference could not be attributed to natural selection, for the differences in height, although intrinsic, would be variable due to the effects of the environment.

Thus, under my modified version of Godfrey-Smith's framework, population differences in intrinsic-invariable properties within an environmental background leading to some differences in reproductive outputs are the ones to be attributed to natural selection. Conversely, differences in

---

[16] I write "usually" because some plants need fire for their seeds to germinate, for instance.

[17] This example is mistakenly attributed to Scriven (1959), who was rather talking of individuals sitting where a bomb or a tree falls.

reproductive outputs due to differences between members of the population in intrinsic-variable and extrinsic properties within an environmental background should be attributed to drift.

Responses to Objections

One worry with the new framework is that it will only be legitimate to talk about natural selection with respect to genes, since arguably only genes are invariable. To this worry, I have two responses. First, the claim that only genes are invariable can be shown to be wrong when considering that other developmental resources such as the centrosome or mitochondria are relatively invariable factors that are furthermore transmitted at each generation (Lu and Bourrat 2016). More generally, epigenetic factors (understood broadly) might be invariable. This means that invariability is not necessarily restricted to molecular genes. Second, having two legs, a particular color of eyes, or any other qualitative trait, might be invariable under a wide range of conditions. But this response should be completed by the response to another worry, which is that intrinsic-invariable properties turn out to be the only ones that are heritable. More specifically, the worry is that intrinsicness and heritability are sufficient to recover natural selection without having to invoke invariability. This worry would be partly justified if heritability were understood strictly as the ratio of the genetic additive component of variance on total phenotypic variance. But the notion of heritability has several meanings and it is not clear which ones of these different meanings is more fundamental (Downes 2009; Bourrat 2015b; Bourrat and Lu 2016; Lynch and Bourrat 2016). When heritability is understood more abstractly than the definition above, it does not guarantee invariability, so heritability and invariability are different at least under some sense of heritability. Second, considering the Breeder's equation mentioned above, heritability explains the response to selection (evolutionary change) rather than selection itself. That is, heritability can be used to give conditions for natural selection the *product* rather than natural selection the *process*. Invariability is about the (physical) process.

Another type of worry, related to the two previous ones, is that the variation over time of many phenotypes is itself the result of natural selection. Examples are pervasive and include sex change in fish, migration schedules, or life history traits. To answer this worry one can remark that although these phenotypes are variable they can nevertheless be associated with invariable phenotypes, namely life history trajectories. Here is one possible example. Suppose, for instance, that two types of organisms grow over time, with different patterns, and that height at every point in time is causally linked to viability (maybe in different ways at some stages of the organisms' lives). Suppose that although height varies over time, and it has some consequences on reproductive output, the developmental trajectory with respect to height is the same (modulo the effects of some differences in extrinsic properties) within each type. In such a case, invariability is recovered at some level (namely the developmental trajectory) in spite of phenotypic variability. If this is correct, one should only consider associating with natural selection intrinsic-variable properties that can also be associated with a particular life trajectory.

Yet another worry is that what looks to be invariable depends on the grain of description one is using. For instance, an individual could be "red" from a coarse grained perspective and remain "red" over time, but be "scarlet" from a fine-grained perspective and change to "crimson" over time. This is a problem ultimately linked to the reference class problem in probability. There is currently no satisfactory solution to this problem, and I will not attempt to provide one here (although see Strevens 2016 for more on this problem in relation to drift).

The final and most serious worry with my framework, on which I will spend more time, is that it seems to leave out one important feature classically associated with drift, namely randomness. Randomness is in fact mentioned in every textbook of population genetics. For instance, Hamilton (2009) defines (genetic) drift as: "*Random* changes in allele frequency from one generation to the next in biological populations due to the finite samples of […] alleles that contribute to the next generation" (pp. 54–55, emphasis added). Drift, following this definition, corresponds thus to the *random* deviations from the expectation of allele contributions between two generations. That is, the deviations are supposed to vary independently from types or alleles. Yet, starting from the physical interpretation of drift in terms of differences in extrinsic (and intrinsic-variable) properties provided above, one can easily imagine cases in which there are differences in extrinsic (or intrinsic-variable) properties between the members of the population but in which those properties are more often associated with one type or allele rather than another (correlation without causation between the intrinsic and extrinsic properties) and thus predictable to some extent.[18]

In such cases, it would be extremely problematic to consider the resulting evolutionary change as "random,"

---

[18] Note here that this problem is not encountered in Godfrey-Smith's initial account. This is because in his account of drift, the parameter C introduces an element of randomness.

since it is predictable.[19] The phenomenon in which some extrinsic properties are independently correlated with one type of entities is a more general case of what Godfrey-Smith (2009, p. 28), following the terminology used in quantitative genetics, calls "correlated response," a shorthand for "correlated response to selection." A pure case of correlated response to selection occurs when one phenotype (neutral with respect to selection) happens to have different frequencies at two different generations merely due to its physical links to another phenotype that is selected (both phenotypes are supposed to belong to the same entity). Linkage disequilibrium is a classical cause for correlated responses to selection. Using Sober's (1984) terminology, a trait that changes in frequency because it is correlated to another one being selected represents a case of selection *of*. The concept of "correlated response"—this time distinct from "correlated response to selection" and not merely a shorthand—can be generalized to any case involving a correlation between an entity's properties and its reproductive output whether or not these properties are intrinsic-invariable. For instance, if some entities (or types) happen to be in a favorable part of the environment for reasons independent from their intrinsic-invariable properties and as a result have more offspring, this qualifies as a correlated response, in which cases properties different from intrinsic-invariable of the entities produce the correlation. In the remainder of the article, by "correlated response" I will refer to this more general notion. Because the notions of differences in extrinsic and intrinsic-variable properties might in some cases lead to phenomena close to, but different from drift, namely correlated responses, it appears that they are insufficient, alone, to fully specify drift,[20] which typically implies the notion of randomness while the correlated response does not.

Luckily, using the framework of philosophers such as Rosenthal (2010), Strevens (2011) or Abrams (2012a, b), who independently developed different, but very similar interpretations of probabilities to understand *deterministic* systems (e.g., wheels of fortune and tosses of coins) in which the probabilities of events have a "propensity-like look and feel" (Strevens 2011), one can separate cases of drift that in some sense are "genuinely" random from cases of correlated response. I put genuinely in scare quotes, because strictly speaking deterministic systems do not exhibit genuine randomness (unless their initial conditions are genuinely random), but they can have a number of properties that makes their outcomes, under some conditions, unpredictable.[21] Although the interpretations of Rosenthal, Strevens, and Abrams are subtly different, for the purpose of this article, I will treat them as similar. Following Rosenthal (2010), I will term this interpretation of probability the "natural-range" interpretation. Note again that by "deterministic system" I am not talking about the inputs of the system that will give the system its initial conditions. These inputs might be deterministic or indeterministic, and they might be known or unknown. Note also that it is this conception of determinism that Rosenthal (2010, pp. 77, 79–80) has in mind.

Going into the details of this interpretation would greatly exceed the scope of this article, but I can very briefly introduce the key ideas. Roughly speaking, this interpretation tells us that the probabilistic pattern of some systems relies on two objective properties of those systems' initial-state space:

> On the one hand, in any (not too) small vicinity of an initial state leading to a given outcome *a* there are initial states leading to different outcomes, which explains why the outcome of a single trial cannot be predicted. On the other hand, for each outcome, the proportion of initial states leading to it is constant all over the initial-state space, i.e., it is approximately the same in any not-too-small segment of the space, which explains why there are certain stable characteristic relative frequencies with which the different outcomes occur. (Rosenthal 2010, p. 75).

If the initial-state space of the system has these two properties, Rosenthal tells us that one can obtain the objective probability of an event by dividing the standard measure of the intersection between a not-too-small equilateral interval of initial states (with n dimensions) and the set of initial states that lead to a given outcome, with the standard measure of the equilateral interval of initial states. In other words, this conception of probability tells us that the probability of an outcome is roughly equal to the proportion of initial states in a given not-too-small interval of initial states that leads to this outcome.

To successfully apply this account to biological populations and for evolutionary theory, to capture "genuinely" random drift, it is thus necessary that biological populations exhibit Rosenthal's two properties. If a particular evolutionary outcome is measured between two generations in a population with two types satisfying these two properties, this means the same outcome could have been obtained however the organisms of the two types were

---

[19] Note that the contrapositive is not true: something unpredictable might not necessarily be random.

[20] Note however that some might want to define drift more broadly than it is classically done and include cases of correlated responses as cases of drift.

[21] Note that by "unpredictable" I am referring here to a single outcome. In the case of drift this would amount to the evolutionary success of a single allele. At the population level it is quite predictable that a population will exhibit drift.

distributed in the environment at the beginning of the first generation, but also that considering a not-too-small region of the initial state space, a different outcome could have been obtained with very similar initial conditions. Drift, when these assumptions are verified, corresponds to the deviation from what would have happened, had the number of members of each type been larger and/or had the number of microstates leading to higher-than-average reproductive output in the population been larger, so that the access for the different types of entities to the different types of microstate would have been the same. Two variations of these cases have been briefly described above.

Although the two assumptions of the natural-range interpretation of probability look reasonable, it seems that in many real biological examples they will not be met. Let us recall that correlated responses will arise if the entities of the population are in a heterogeneous environment and that different entities (types), for reasons independent from their intrinsic-invariable properties, have different access to different parts of the environment. In such cases, it seems clear that the assumptions that the same evolutionary outcome would be obtained however the two types were distributed, or that any possible evolutionary outcome would be obtained in very similar initial conditions, will not be tenable.

Imagine, for instance, two types of dandelion, *A* and *B,* with equal fitness, assuming a similar setup as the one used in the second section. The seeds of type *A* more often land on a favorable patch than seeds of type *B,* for reasons independent from their intrinsic-invariable properties at the grain of description used, because more favorable patches happen to be those closer to a river. If at the second generation type *A* has increased in frequency, it is hard to imagine that it would also have happened had *A* grown further away from the river and had *B* grown closer to the river. Similarly it is hard to conceive that had the conditions been almost identical *A* would have decreased in frequency. In fact, the river being on average closer for seeds of type *A* than of type *B*, the seeds of *A* would still have had higher chance to land in the favorable patches. Thus, in this case Rosenthal's two conditions seem to be violated.

One possible response to this specific problem could be that if the different entities of a set of entities have a differential access to different parts of the habitat, and this leads to differences in reproductive outputs, it might be justified to consider this set of entities not as one but two populations. But this solution will only be satisfactory in the limit case in which the environment is highly discrete (or nearly so), that is, distributed in patches, and where each patch is sufficiently sized to carry a large enough number of entities and with limited migration between the patches so that it is reasonable to call the entities on a patch "a population." The risk otherwise is to end up with a very large number of populations of only one individual or a population in which migration is overwhelmingly high. Many simple cases of heterogeneous environment are continuous rather than discrete. For instance, the simple case of the dandelion seeds and the river above is a case of continuous heterogeneity (although for simplicity I have presented it in a discrete fashion), since the distance to the river is a continuous parameter. This renders this solution difficult to implement.

Another solution is to consider that *A* and *B* being on average at different distances from the river, not to be the relevant initial conditions for this case, and that if one was considering the relevant initial conditions to be at an earlier time, it could have happened that with very similar initial conditions *B* and not *A* ends up closer to the river, while at the same time *A* could have also been closer to the river with very different initial conditions. This is one solution proposed by Rosenthal (2010) for the general case. I am not convinced that it can be applied straightforwardly in evolutionary biology. Note that I am not contending here about the theoretical validity of this solution, but merely pointing out the difficulty to implement it in some biological contexts. One aim in population genetics is to understand whether and to what extent drift and natural selection are responsible for the evolutionary change one can observe across generations. If one knew that *A* and *B* were on average at different distances from the river in a given population, this fact would typically not be considered as being the result of some earlier initial conditions. Rather, distance from the river would be considered as part of the initial-state space in this population, and predictions about evolutionary change would be made from this initial-state space that does not satisfy Rosenthal's two conditions. Perhaps in some contexts of studying evolutionary changes over longer time scales than a few generations and across a large number of populations of dandelions, distance from the river could be considered as an outcome of earlier initial conditions of a system that satisfies Rosenthal's two conditions. Yet this solution cannot serve as a general one.

Leaving for further work the different problems of applying the natural-range interpretation to evolutionary theory (including drift), suffice it to say that the classical notion of drift, that is, the one involving unpredictability of evolutionary change and relying on the notion of expected value, can be approximated in deterministic populations exhibiting similar outcomes for different initial conditions and all the possible outcomes for very similar initial conditions. Drift under this interpretation corresponds to cases in which the different types of entities forming a population have a different access to the types of environmental microstates possible in this population due ultimately to differences in extrinsic properties and with downstream consequences in terms of reproductive outputs. In populations that do not exhibit these properties, depending on whether one is ready to accept the sketches of solutions above, it is less clear whether differences in extrinsic (and intrinsic-variable) properties can be strictly associated with the classical notion of drift, if drift is thought to be conceptually distinct from cases of correlated responses. At any rate, a differential access to microstates of the environment between the different types of entities of a population due to differences in extrinsic properties that lead to differences in reproductive output should not be considered as cases of natural selection, but as resulting from a different evolutionary process. One might thus want to separate this general notion of drift relying on extrinsic properties and access to microstates into two subtypes one could call *natural-range drift,* from which the classical probabilistic concept of drift can be recovered, and *broad-sense drift,* which will include the cases of correlated response.

Conclusion

I have proposed a physical interpretation of drift that is different and complementary to the interpretation of drift borne out of the propensity interpretation of fitness and that satisfies the two desiderata outlined in the introduction. I have shown how deviations from expected reproductive output could be interpreted as arising from differences in extrinsic and intrinsic-variable properties (at a given grain of description), using some fictive examples of biological populations. My account is also "agnostic" about the initial conditions of the system, which might be deterministic or not.

Acknowledgments I am thankful to Mark Colyvan, Paul Griffiths, Aidan Lyon, Maureen O'Malley, Charles Pence, Arnaud Pocheville, and Kim Sterelny for their comments on a previous version of the manuscript. This research was supported under the Australian Research Council's Discovery Projects funding scheme (Project Number DP150102875).

References

Abrams M (2009) The unity of fitness. Philos Sci 76(5):750–761

Abrams M (2012a) Mechanistic probability. Synthese 187(2):343–375

Abrams M (2012b) Mechanistic social probability: how individual choices and varying circumstances produce stable social patterns. In: Kincaid H (ed) The Oxford handbook of philosophy of social science. Oxford University Press, Oxford, pp 184–226

Abrams M (2014) Environmental grain, organism fitness, and type fitness. In: Barker G, Desjardins E, Pearce T (eds) Entangled life. Springer, Netherlands, pp 127–151

Beatty JH, Finsen SK (1989) Rethinking the propensity interpretation: a peek inside Pandora's box. In: Ruse M (ed) What the philosophy of biology is: essays dedicated to David Hull. Kluwer Publishers, Dordrecht, pp 17–30

Bouchard F, Rosenberg A (2004) Fitness, probability and the principles of natural selection. Br J Philos Sci 55(4):693–712

Bourrat P (2015a) Distinguishing natural selection from other evolutionary processes in the evolution of altruism. Biol Theory 10(4):1–11. doi:10.1007/s13752-015-0210-6

Bourrat P (2015b) How to read "heritability" in the recipe approach to natural selection. Br J Philos Sci 66(4):883–903

Bourrat P, Lu Q (2016) Dissolving the missing heritability problem. Philos Sci

Brandon RN (1978) Adaptation and evolutionary theory. Stud Hist Philos Sci Part A 9(3):181–206

Brandon RN (1990) Adaptation and environment. Princeton University Press, Princeton

Brandon RN, Carson S (1996) The indeterministic character of evolutionary theory: no 'no hidden variables proof' but no room for determinism either. Philos Sci 63(3):315–337

Downes SM (2009) Moving past the levels of selection debates. Biol Philos 24(5):703–709

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

Falconer DS (1981) Introduction to quantitative genetics. Longman, New York

Gillespie JH (2004) Population genetics: a concise guide. Johns Hopkins University Press, Baltimore

Glymour B (2001) Selection, indeterminism, and evolutionary theory. Philos Sci 68(4):518–535

Godfrey-Smith P (2007) Conditions for evolution by natural selection. J Philos 104(10):489–516

Godfrey-Smith P (2009) Darwinian populations and natural selection. Oxford University Press, New York

Grafen A (2000) Developments of the Price equation and natural selection under uncertainty. Proc R Soc Lond B Biol Sci 267(1449):1223–1227

Hájek A (2012) Interpretations of probability. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/archives/win2012/entries/probability-interpret/

Hamilton M (2009) Population genetics. Wiley-Blackwell, Oxford

Lu Q, Bourrat P (2016) The evolutionary gene and the extended evolutionary synthesis. Br J Philos Sci

Lynch KE, Bourrat P (2016) Interpreting heritability causally. Philos Sci. doi:10.1086/688933

Mills SK, Beatty JH (1979) The propensity interpretation of fitness. Philos Sci 46(2):263–286

Millstein RL (2002) Are random drift and natural selection conceptually distinct? Biol Philos 17(1):33–53. doi:10.1023/a:1012990800358

Millstein RL (2003) Interpretations of probability in evolutionary theory. Philos Sci 70(5):1317–1328. doi:10.1086/377410

Millstein RL (2016) Genetic rift. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Fall 2016. http://plato.stanford.edu/archives/fall2016/entries/genetic-drift/

Okasha S (2006) Evolution and the levels of selection. Oxford University Press, New York

Pence CH, Ramsey G (2013) A new foundation for the propensity interpretation of fitness. Br J Philos Sci 64(4):851–881

Pfeifer J (2005) Why selection and drift might be distinct. Philos Sci 72(5):1135–1145

Plutynski A (2007) Drift: a historical and conceptual overview. Biol Theory 2(2):156–167. doi:10.1162/biot.2007.2.2.156

Ramsey G (2006) Block fitness. Stud Hist Philos Sci Part C 37(3):484–498

Ramsey G (2013) Driftability. Synthese 190(17):3909–3928. doi:10.1007/s11229-012-0232-6

Rice SH (2004) Evolutionary theory: mathematical and conceptual foundations. Sinauer, Sunderland

Rice SH (2008) A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. BMC Evol Biol 8:262. doi:10.1186/1471-2148-8-262

Rosenberg A (2001) Discussion note: indeterminism, probability, and randomness in evolutionary theory. Philos Sci 68(4):536–544

Rosenthal J (2010) The natural-range conception of probability. In: Ernst G, Hüttemann A (eds) Time, chance, and reduction: philosophical aspects of statistical mechanics. Cambridge University Press, Cambridge, pp 71–90

Scriven M (1959) Explanation and prediction in evolutionary theory. Science 130:477–482

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

Sober E (2001) The two faces of fitness. In: Singh RS, Krimbas CB, Paul DB, Beatty J (eds) Thinking about evolution: historical, philosophical, and political perspectives. Cambridge University Press, Cambridge

Sober E (2010) Evolutionary theory and the reality of macro-probabilities. In: Eels E, Fetzer JH (eds) The place of probability in science. Springer, Dordrecht, pp 133–161

Stamos DN (2001) Quantum indeterminism and evolutionary biology. Philos Sci 68(2):164–184

Stephens C (2004) Selection, drift, and the ‘forces’ of evolution. Philos Sci 71(4):550–570. doi:10.1086/423751

Strevens M (2011) Probability out of determinism. In: Beisbart Hartmann (ed) Probabilities in physics. Oxford Univeristy Press, New York, pp 339–364

Strevens M (2016) The reference class problem in evolutionary biology: distinguishing selection from drift. In: Ramsey G, Pence CH (eds) Chance in Evolution. University of Chicago Press, Chicago, pp 145–175

Weatherson B (2014) Intrinsic vs. extrinsic properties. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/archives/fall2014/entries/intrinsic-extrinsic/

Weber M (2001) Determinism, realism, and probability in evolutionary theory. Philos Sci 68(3):S213–S224