Acta Biotheor (2018) 66:159–176
https://doi.org/10.1007/s10441-018-9331-1

REGULAR ARTICLE

Natural Selection and Drift as Individual-Level Causes of Evolution

Pierrick Bourrat[1,2]

Received: 2 May 2018 / Accepted: 7 May 2018 / Published online: 14 May 2018

Abstract In this paper I critically evaluate Reisman and Forber's (Philos Sci 72(5):1113–1123, 2005) arguments that drift and natural selection are population-level causes of evolution based on what they call the manipulation condition. Although I agree that this condition is an important step for identifying causes for evolutionary change, it is insufficient. Following Woodward, I argue that the invariance of a relationship is another crucial parameter to take into consideration for causal explanations. Starting from Reisman and Forber's example on drift and after having briefly presented the criterion of invariance, I show that once both the manipulation condition and the criterion of invariance are taken into account, drift, in this example, should better be understood as an individual-level rather than a population-level cause. Later, I concede that it is legitimate to interpret natural selection and drift as population-level causes when they rely on genuinely indeterministic events and some cases of frequency-dependent selection.

Keywords Natural selection · Drift · Evolution · Causality · Probabilities · Manipulation · Invariance

✉ Pierrick Bourrat
p.bourrat@gmail.com

[1] Department of Philosophy, Macquarie University, North Ryde, NSW 2109, Australia

[2] Department of Philosophy, School of History and Philosophy of Science and Charles Perkins Centre, The University of Sydney, Sydney, NSW 2006, Australia

1 Introduction

Population geneticists classically distinguish four causes[1] of evolution: natural selection, mutation, migration and drift.[2] Yet, what an "evolutionary cause" is, is not immediately clear. It has been and still is subject to an important debate in philosophy of biology.[3] Most of the literature focuses on natural selection and drift while leaving out mutation and migration. I follow suit.

The philosophical landscape on the causes for evolutionary change can roughly be divided into three views. First, some philosophers think it is misleading to consider natural selection and drift as causes. According to them, what we call natural selection is a mathematical aggregate of unique events happening to individuals forming populations (e.g. Matthen and Ariew 2002; Walsh et al. 2002). Following the literature, I call this view the *statistical view*. Under this view, each individual trait (or, depending on the particular account, type of organism) in a population is assigned a *trait-fitness* value—in its most abstract form an expected (mathematical) growth rate—which is the result of different events occurring at the individual-organism level such as death, birth or mating. In that sense, statisticalists locate causation at the individual rather than the population level. Yet, those events, they claim, cannot be equated with natural selection (or drift). This is because trait-fitness is a non-causal (statistical) property of types at the population level. Consequently, differences in trait-fitness in a population entail evolution by natural selection mathematically (statistically) rather than causally (Matthen and Ariew 2009).

Over the years, the statisticalists have nuanced some of their claims. For instance, Ariew et al. (2015) defend statisticalism on the ground that it provides autonomous-statistical explanations, that is explanations which do not cite causes, but only properties of probability distributions. The point here is not that natural selection and drift are not causes of evolutionary change. Rather, it is that one *can* talk about natural selection and drift without having to refer to causality.[4] Walsh et al. (2017) (see also Walsh 2007, 2010) defend statisticalism on the ground that explanations involving natural selection and drift are relative to a model. According to them, this weakens the idea of viewing natural selection and drift as causes, since causal

[1] In population genetics, the term 'force' is used more often than the term 'cause.' I consider them as roughly equivalent, although I will use the term 'cause' throughout the paper. This is because I will make use of concepts developed in the philosophical literature on causation. For more on the difference between the use of these two terms in evolutionary theory see Hitchcock and Velasco (2014).

[2] This division into four causes is to some extent arbitrary. Some authors would for example add sex or recombination as other forces or causes.

[3] See for example Bouchard and Rosenberg (2004), Matthen and Ariew (2002, 2009), Millstein (2006), Northcott (2010), Reisman and Forber (2005), Rosenberg and Bouchard (2005), Shapiro and Sober (2007), Stephens (2004), Walsh (2000, 2007, 2010), Walsh et al. (2002), Ariew et al. (2015), Walsh et al. (2017), and Otsuka (2016).

[4] To be sure, most causal explanations involve the use of statistics in evolutionary biology. The statisticalists have no objection to that, but they claim that referring to causes is in principle dispensable for evolutionary explanations.

relations are commonly regarded as "objective, description-independent features of the world."[5]

Second, Bouchard and Rosenberg (2004; see also Rosenberg and Bouchard 2005; Otsuka 2016) offer a very different view, in direct opposition to that of the statisticalists. Not only do they argue that natural selection is a causal process, they also claim that fitness should not be understood solely as a population-level statistical property. They conceive of fitness as a context-dependent property of *individual*[6] *entities* forming populations. This property must be distinguished from a growth rate or reproductive output, since Bouchard and Rosenberg view a difference in growth rate between two types of entities not as a difference in fitness, but as a *consequence* of a difference in fitness. In other words, growth rates are proxies for fitness (a proxy for fitness is also called *realized* fitness). As Bouchard and Rosenberg (2004, p. 710) put it: "selection [is] a contingent causal process in which *individual* fitnesses are the causes and subsequent *population* differences are the effects" (my emphases). Thus, under Bouchard and Rosenberg's view only causes of evolutionary change are at the individual level, not the consequences of these causes. I call Bouchard and Rosenberg's view the *individual-cause view*.

A number of authors have proposed a third view, namely that natural selection and other evolutionary processes are causes of evolutionary change, not at the individual level, as argued by Bouchard and Rosenberg, but at the population level (e.g. Millstein 2006; Reisman and Forber 2005; Stephens 2004). One argument from the *population-cause view* proponents is that since natural selection and drift can systematically be manipulated at the population level (e.g., by changing population size we can change the influence of drift), they represent genuine causes at that level under some legitimate account of causation. This is what Reisman and Forber (2005) call the 'manipulation condition,' following Woodward's (2003, 2013) now popular manipulationist account of causation. Manipulating (also known as an 'intervention on') a variable X to establish whether it is a cause of a second variable Y, consists, following Woodward (2003, p. 14), in changing "the value of X in such a way that if any change occurs in Y, it occurs only as a result of the change in the value of X and not from some other source."

One important point to note here is that both the individual-cause and the population-cause proponents agree that the evolutionary consequences of natural selection are at the population level. Their disagreement lies in whether the causes of natural selection are population-level or individual-level causes.

In this paper, I do not engage in the disagreements between the *statisticalists* and the *causalists* (represented both by the individual-cause and population-cause

[5] Otsuka et al. (2011, see also Northcott 2010), argue that context independence is far from being a necessary condition for a relationship to be considered as causal. Context dependence is in fact a feature of causal modeling.

[6] In this paper, "individual" should be understood here as "any entity below the level of the population." As pointed out by Charles Pence (personal communication), this notion of "individual" might refer to something different from what the individual causalists mean by this term, since for them "individuals" often refer to "individual organisms." If that is the case, I accept this departure from that of the individual causalists.

views). For an excellent analysis of the debate that opposes them see Otsuka (2016) who, starting from a causal-modeling perspective—a mathematical account of the manipulationist account of causation (Pearl 2009)—defends the individual-cause view against the statisticalists. Otsuka convincingly argues that the statisticalist position is untenable on a number of grounds and a rebuttal of his arguments is yet to be provided by the statisticalist camp. Delving into the details of the arguments would greatly go beyond the scope of this paper. The interested reader can refer to Otsuka's article.

My focus will rather concern one aspect of the opposition between the individual-cause and the population-cause view which has received much less attention in the literature (for an exception see Millstein 2006, especially Sect. 4.2). I argue that the manipulation condition proposed by Reisman and Forber is, in and of itself, insufficient to favor the population-cause view against the individual-cause view. To illustrate their point, Reisman and Forber use a controlled experiment realized by Dobzhansky and Pavlovsky (1957) which, according to them, demonstrates that one form of drift, namely the founder effect, is a population-level cause.

I show that although the population-level cause view is available and standardly invoked in population genetics, an individual-level causal explanation is also available in cases of frequency-independent selection. I argue that this individual-level explanation is superior to the population-level one. As I show, this is precisely the case for the phenomenon described by Dobzhansky and Pavlovsky and used by Reisman and Forber in their analysis. The reason I identify it as superior is that the causal relationship in the individual-level explanation is more *invariant under intervention*, following Woodward's (2000, 2003, 2010) concept of invariance, than the one identified with the population-level explanation. Invariance under intervention measures the extent to which a relationship between two variables satisfying the manipulation condition remains stable or unchanged as various other changes are made in the background of this relationship.

My arguments thus demonstrate that drift (but the same can be argued for natural selection), at least in some cases, should be considered as an individual rather than a population-level cause. That said, in the last section I show that in indeterministic setups and in cases of frequency-dependence selection it can be legitimate to consider drift and natural selection as population-level causes. I briefly relate this point to a defense of the population-cause view proposed by Millstein (2006).[7]

[7] Stephens' (2004) position seems close to that of Reisman and Forber. He assumes that drift and NS are population-level causes: "the point is that the effect of drift is only properly understood at the population level. It is a population level cause. One sees the differential causal impact of drift only by comparing populations of different sizes" (p. 556, see also pp. 563–564).

2 The Manipulation Condition

Reisman and Forber's (2005) main argument for arguing that natural selection and drift are causes of evolutionary change resides in what they call the 'manipulation condition', which can be formulated as follows:

> (Manipulation Condition) If a variable *A* can systematically be manipulated and bring about changes in variable *B*, then *A* is a cause of *B*.

This condition is the most important feature of Woodward's manipulationist account of causation (Woodward 2003, 2013), which is the account of causation Reisman and Forber rely on to make their claims and that I will adopt throughout the paper. From there, Reisman and Forber develop a simple deductive argument which I reconstruct as follows:

> [Premise 1] Manipulation Condition.
>
> [Premise 2] Manipulating drift and natural selection in a systematic way using population-level variables can bring about changes in the evolutionary dynamics of a population.
>
> [Conclusion] Natural selection and drift are population-level causes of evolutionary change.[8]

Before going further two points should be noted. First, my formulation of Reisman and Forber's argument is different from their original formulation. Although Reisman and Forber argue that natural selection and drift are population-level causes, they do not make it explicit until they defend their position against the *statistical* view, which like the *individual-cause* view locates causation at the individual level but disagrees that fitness is an individual property. Reisman and Forber's initial argument is only that the manipulation condition can be used to argue for natural selection and drift as being *causes* of evolutionary change, without specifying the level. Later they argue that manipulating population variables (like population size in the case of drift) leads to change in evolutionary outcomes (for specific quotes see footnote 8), and that these population variables are causes of evolutionary change. Thus, to make clear the point that Reisman and Forber apply the manipulation condition at the population rather than the individual level, I have modified their argument accordingly.

Second, in a follow up paper, Forber and Reisman (2007) claim that there is no privileged level of analysis at which *all* the causal relationships are located. That is to say, lower levels of organization should not necessarily be regarded

[8] Reisman and Forber's exact formulation is as follows: "[Premise] 1. The manipulation condition. [Premise] 2. Manipulating the character of selection and drift can result in systematic changes to population-level dynamics. [Conclusion] Selection and drift are causes of population-level dynamics." (2005, 1114) Reisman and Forber explicitly refer to population-level variables as causes when they claim: "we can manipulate the strength of drift in a population by manipulating the size of the population" (p. 1115). When they refer to the authors of the experiment they discuss, they write "[t]hey manipulate a population-level parameter to test how selection and drift interact to produce evolutionary change" (p. 1116).

as having a privileged role in explanations. The manipulation condition can be applied whether variables are defined at the individual or at the population level. I am very much in agreement with this claim since this is one of the desirable features of Woodward's (2003) account of causation. My contention with Reisman and Forber is that in the case they present (and similar ones) where both a manipulation can be made at the population level and at the individual level, the individual level, for reasons developed in Sect. 4, is a better level of analysis.

Starting from one experiment (see Sect. 3), Reisman and Forber (2005, pp. 1120–1122) offer three arguments to defend the population-cause view against both individualist positions (statistical and individual-level). The first is that they succeed in showing that drift can be "manipulated" at the population level without any reference to individual-level properties. Although the main example they use concerns a particular case of drift, namely the founder effect, similar demonstrations using only population-level variables for other forms of drift and for natural selection could be made (Reisman and Forber, 1119). The founder effect is defined by Futuyma (2005, p. 548) as "[t]he principle that the founders of a new population carry only a fraction of the total genetic variation in the source population." Drift, more generally, is difficult to define succinctly as it is associated with a diversity of phenomena (Beatty 1984; Millstein 2016; Plutynski 2007). Beatty provides a definition of drift as phenomena that "have one or another biological form of random or indiscriminate sampling, and consequent sampling error" (1984, p. 273). In the case of the founder effect, the sampling error originates from sampling particular alleles in the founder population that are unrepresentative of the frequencies of alleles found in the larger population. Second, Reisman and Forber argue that even if population-level properties supervene on individual-level properties they do not necessarily reduce to them. Third, they argue that speaking of natural selection and drift in terms of population-level causes is more informative than speaking about them in terms of individual-level causes.

I see no problem with Reisman and Forber's first claim that drift can be manipulated at the population level without reference to the individual level. And in fact, this is routinely how population geneticists reason about drift, where it is estimated in part from the population size (more precisely the effective population size) (Hamilton 2009, Chapter 3). However, I argue below that Reisman and Forber, when illustrating their argument, have succeeded in establishing neither that causes of evolutionary change are population-level causes *irreducible* to individual-level causes, nor that the population-cause view is superior to the individual-cause view. On the contrary, I show that an explanation using only individual-level properties is available and that it is superior to that of Reisman and Forber, at least for the example they present. To be fair to Reisman and Forber, as already pointed out, the target of their paper is mostly the statistical view rather than the individual-cause view. Yet, many of their claims apply to both views since they both consider the individual level as a privileged one for causal explanations of evolutionary change.

3 Dobzhansky and Pavlovsky's Experiment Reconsidered

To illustrate their point, Reisman and Forber present a controlled experiment realized by Dobzhansky and Pavlovsky (1957) which shows that the founder effect can be manipulated at the population level and can cause deviations from expected evolutionary change. "Expected evolutionary change" refers here to the evolutionary change one would observe if natural selection was the only cause of change, that is if the fitnesses of the individuals forming the population were expected fitnesses. Okasha (2006, p. 32) defines expected fitness as "the number of offspring the entity would on average produce if it found itself in the same environment repeatedly." Since Dobzhansky and Pavlovsky focus on the founder effect, by *drift* I will hereafter mean the *founder effect*, which is the target of Reisman and Forber's argument, but it should be clear that I intend the demonstration to be applicable to other phenomena (but not all phenomena, see Sect. 5) typically associated with drift and natural selection. Drift, in this experiment, is manipulated through initial population size.

The experiment shows that the frequency of a particular allele (initially at a frequency of 0.5) at a locus with two alleles, in different replicates of populations sampled from a source population of fruit flies exhibits more variability over time when the populations sampled are initially small (20 fruit flies) than when they are large (4000 fruit flies). Indeed, when the frequency of the allele of the sampled populations is tracked over time, the variability of the different replicates in their deviation from the (estimated) expected frequency of this allele at equilibrium is higher when the population is composed of 20 fruit flies than when the replicate populations are founded with 4000 fruit flies. Note that Dobzhansky and Pavlovsky refer to populations variants of the third chromosome, that is chromosomes with different gene arrangements (labelled *AR* and *P*), not alleles. But in their experiment these chromosomal variants behave like alleles. To go beyond Dobzhansky and Pavlovsky's experiment, I will consider that chromosomal variants in Dobzhansky and Pavlovsky's experiment are alleles and will call them *A* and *B*, instead of *AR* and *PP* to be more general.

Supposing a homogeneous environment, the explanation given by Dobzhansky and Pavlovsky of their results is that when 20 fruit flies are sampled from the source population the sample contains less genetic variability at other loci than the one they study than when 4000 fruit flies are sampled. How much genetic variability at other loci contains a given sample will typically itself depend on the initial sample size. A homogenous environment implies in this context that the environment of each fruit fly can be considered identical.[9] Dobzhansky and Pavlovsky also assume that the larger the genetic variability at different loci in the population source is, the larger the difference between the two samples will be. Importantly, they stress that only the population sizes of the initial samples of the two conditions are different

[9] Note that there exists some heterogeneity in the genetic background of fruit flies. Thus, the environment is not homogenous from the point of view of the alleles of the population. This difference is crucial for the rest of the argument.

(p. 315 see also the long quote below), so that when deviations from expectations are measured, the two types of samples have essentially the same population size. Thus, Reisman and Forber argue, drift is a function of the size of the initial sample. When the founder population size is small, its representativity of the source population is smaller than when the founder population is large. This, as a result, leads to larger deviations from the expected evolutionary change. Although this (standard) population-level explanation is perfectly legitimate, one should note that the following individual-level explanation[10] is also perfectly compatible with Dobzhansky and Pavlovsky's findings.

In an initially small population (20 fruit flies), the probability that for a given token allele of one type at any generation, another token allele of the other type with the *same* genetic background (or a similar background or again a background with the same or similar effects on fitness) is present in the population is *smaller* than when the sample is large (4000 fruit flies). As a result, when founding population size decreases, there is a higher probability that the cumulated effects on the reproductive outputs of all token alleles *A* by alleles in their background differ from that of token alleles *B*, and that any difference observed is higher than when the population is larger. Consequently, it becomes more probable that the evolutionary outcomes in initially small populations differ from their expectations than in initially large populations. Following this reasoning, manipulating the genetic background of a token allele (an individual-level property), without reference to population size, can thus have some effect on the deviation from an expected evolutionary outcome. Since a deviation from an expected evolutionary outcome is what constitutes drift, this demonstrates that drift can be manipulated without reference to population size.

Before proceeding further, I should respond to one possible objection to the alternative explanation I have just proposed. The objection is that variability in genetic backgrounds does not amount to drift, but rather to selection. In fact, it is well known, following Fisher's fundamental theorem (Fisher 1930; Okasha 2008), that the strength of selection in a population is proportional to the genetic variance in fitness in the population. By homogenizing the genetic backgrounds of fruit flies, one might argue, that it just reduces the variance in fitness of the fruit flies. This objection would be a valid one if the expected frequencies measured in Dobzhansky and Pavlovsky's experiment were referring to the genotype of fruit flies *across their whole genome*. Yet, in their experiment, they only refer to two alleles (*AR* and *PP*). Similarly, Reisman and Forber refer to two types in the experiments. As mentioned in footnote 10, the individuals in Dobzhansky and Pavlovsky's experiment (and population genetics more generally) are alleles. This means that, from the point of view of a token allele, variation in the genetic background of this alleles should be considered as part of its environment (Haig 2012; Okasha 2008: Lu and Bourrat 2017; Sterelny and Kitcher 1988). As expressed by Haig (2012), one of the main proponents of gene selectionism which is largely inspired from classical population genetics:

[10] In the example proposed by Dobzhansky and Pavlovsky, the measures of frequency are made on *alleles*, not individual organisms. Thus, an "individual" refers here to a token allele.

> A gene's environment encompasses all factors that are shared with the alternative against which the gene's effects are measured. It contains not only factors external to the cells and bodies of organisms, but also (and more immediately) these cells and bodies themselves. A body can be viewed as the collectively-constructed niche of the genes of which it is an extended phenotype. Among the most important parts of a gene's environment are the other molecules with which it interacts. *Other genes, even other alleles at the same locus, are parts of a gene's social environment* […]. (p. 465, my emphasis)

The individual-level explanation I provided is perfectly in line with Dobzhansky and Pavlovsky's interpretation of the larger deviations from expectation in small populations when they write: "The segments of the gene pool which arise from race hybridization are smaller, and therefore *less uniform*, in the populations descended from small than in those descended from large numbers of founders" (p. 318, my emphasis). This explanation is equivalent to say that from the perspective of one focal token allele, say *A* in a given small founder population, its genetic background is more likely to be different from that of a token allele of the different type (*B*), than when the founder population is large. This, once again, is vindicated by Dobzhansky and Pavlovsky who explicitly refer to the link between variability of genetic background and drift when they write in reference to populations made of mixed geographic origins (which is the case in their experiment) that:

> [T]he selective fates of the chromosomal gene arrangements become dependent upon the polygenic genetic background, which is highly complex and variable because of the gene recombination that is bound to occur in populations descended from race hybrids. *Here random drift becomes operative and important.* It becomes important despite the populations being small only at the beginning of the experiments, because the foundation stocks in some populations consisted of small numbers of individuals. Thereafter, all the populations expand to equal sizes, fluctuating roughly between 1000 and 4000 adult individuals. Such populations can be regarded as small only in relation to the number of gene recombinations which are possible in populations of hybrid origin. (p. 316, my emphasis)

This reasoning also points to the fact that the probability to find identical or similar genetic backgrounds for the two alleles is lower in small rather than large founding populations, only under some assumptions about the population. This is ultimately the reason why drift is commensurate to effective population size and not population size per se (for more on this point see Bourrat 2017), for an *effective* population size is defined as "the size of an ideal Wright–Fisher population that maintains as much genetic variation or experiences as much genetic drift as an actual population regardless of census size."(Hamilton 2009, p. 73)

To make the individual-level explanation more concrete and generalize beyond Dobzhansky and Pavlovsky's experiment, imagine an extreme case in which each fruit fly is haploid and only has two loci. Call one locus the "focal locus", while the other is the "background locus". Suppose there are two alleles at the focal locus (*A* and *B*) and 10 at the background locus, each with a different effect on the

reproductive output of the fruit flies. We assume that each allele at the background locus has the same frequency in the global population (0.1), that is the population from which the small and larger founder samples are made. Finally, suppose that each one of the twelve alleles has a different effect on the reproductive output of fruit flies. Under these assumptions, in a sample with a founder population of 20 fruit flies (10 *A*, 10 *B*), everything else being equal, it is less likely that for any token allele of a given type (say *A*) with a given allele at the background locus, there is another token allele of the other type (*B*) in the sample with the same allele at the background locus than when the sample is made of 4000 flies (2000 *A*, 2000 *B*).

The probabilities can indeed be calculated using a binomial distribution. In the small founder population the probability that for a fruit fly *A* there is at least one fruit fly *B* with the same background is given by the cumulative binomial probability of at least one success in 10 independent trials, with the probability of success being 0.1.[11] In fact there are 10 types of background alleles with the same frequency in the population, only one of which will be same as the background allele of the particular fruit fly chosen (the success). This probability is approximately 0.65. If we now calculate the same binomial probability but the number of trials is 2000 (large initial founder population), the probability becomes almost 1. The difference between these two probabilities demonstrates that, when there is variation on the background locus, it is more likely that the frequency of the two alleles diverge from their expectation in a small founder population than in a large one originating from the same source population.

Yet, as mentioned earlier, sample population size matters here only insofar as it decreases the probability that for each token allele of one type there is another token allele of the other type with the same genetic background. This probability could in principle be manipulated by intervening directly on the genetic background of fruit flies. A manipulation that would lead the genetic background of the token allele (say *A*) increasing the total variability of the genetic background of the population would lead to an increase in the deviation from the expected frequency of *A*, whatever the sample population size is.[12] On the contrary the manipulation that would decrease the variability would lead to a decrease in deviation from the expected frequency of *A*.

By simply considering the genetic background of one allele to be the genetic environment of this allele, *pace* Reisman and Forber, a population-cause explanation is perfectly reducible (in principle) to an individual-level one. In my example, the variable manipulated is not population size but token alleles' genetic backgrounds. This thought experiment vindicates that manipulating the founder population size

[11] Note, that using the binomial distribution implies the assumption of a founding population of unlimited size (equivalent to a drawing without replacement). If we were to release this assumption and have a source population with a finite size (equivalent to a drawing with replacement), the appropriate distribution would be a hypergeometric one. It can be shown that as the population size increases, the hypergeometric distribution tends toward a binomial one.

[12] This assumes that the variability in the genetic background is not neutral. Of course, such a manipulation would have consequence of a higher magnitude in a small than large population since the frequency of one allele is a small population is different from that of a large one. But that is beside the point.

with no control on the diversity of the genetic background of fruit flies or manipulating the diversity of the genetic background *independently from* the founder population size are equivalent manipulations that can be performed at the same phase of the experiment, more particularly at the generation F2 if referring to Dobzhansky and Pavlovsky's experiment.

4 The Criterion of Invariance

So far, Reisman and Forber's argument is not undermined. I have merely shown that besides a population-level explanation of the founder effect, there is a competing explanation at the individual-level of the same phenomenon. In all conscience, I do not imply that this explanation has been totally disregarded by Reisman and Forber when they write: "[t]he total amount of genetic variation in the replicate populations should affect the outcome of subsequent selection on the populations" (p. 1116). Yet, what they seem to disregard is the fact that this explanation can be cashed out from an individual-level perspective by noticing that manipulating the genetic background a single token allele (an individual-level variable) would affect the deviation from the expected frequency of the allele type.

From there, it is legitimate to ask whether one of the two explanations is superior to the other. Some, at that point, might want to claim that Reisman and Forber's explanation is superior to mine because it has the same explanatory power at a lesser cost: tediously establishing all individual causes versus readily manipulating the founding population size. I argue below, however, that the explanatory power of the individual-level explanation is superior to that of the population-level explanation.

I recall that according to Woodward (2000, 2003, 2010), to be explanatory, a generalization describing a relationship between two or more variables must be invariant where "invariant" (also called "stable") means that "it would continue to hold—would remain stable or unchanged—as various other conditions change" (2000, p. 205).[13] As Woodward stresses, invariance is not an absolute concept. Yet, in evaluating which one of two explanations is the best, invariance will be an important criterion: the more invariant under intervention, the better the explanation: "other things being equal, relationships that are more invariant (and hence more useful for purposes of manipulation and control) provide better explanations" (2003, p. 243). I call this the *criterion of invariance.* Another way to characterize invariance is in terms of context dependence. The less context-dependent a relationship is, the better the causal explanation.[14]

[13] For more on the subtleties of the notion of invariance see Pocheville et al (2017).

[14] This point might look, at first glance, similar to the point made by the statisticalists and discussed in the introduction, that natural selection and drift are not causes of evolution because they are context dependent. The point here however is different. All causal relationships are to some extent context dependent, and following the criterion of invariance, the less context dependent relationship leads to better causal explanation. Thus, the point is *not* that context dependent relationships are not causal explanations as argued by the statisticalists.

Although Reisman and Forber's conceptual apparatus is much less constraining than mine, I show below that their account is nevertheless, following Woodward's criterion of invariance, less explanatory than mine. As a result, it misidentifies drift in terms of population size while ultimately population size is confounded with some individual-level variables that can fully and more invariantly explain the founder effect. In other words, given a population, individual-level variables screen off population-level properties.[15]

One dramatic way of demonstrating this misidentification of drift in terms of population variables is to show that one can *eliminate entirely* the founder effect and drift resulting from it, by manipulating individual-level variables while keeping population size constant. That such a manipulation exists represents a decisive demonstration that population size is causally efficacious only insofar as it depends or supervenes on individual-level variables.

To make this demonstration, imagine that Dobzhansky and Pavlovsky in their experiment, instead of fruit flies (F0) with different genetic makeups, had used fruit flies which were all clones except for the focal locus, in which one sex would have been *AA* while the other would have been *BB*, to produce the generation F1. Suppose also that the fruit flies (F0) had been homozygous at each locus and were reproducing panmictically. Crossing the two types, as in Dobzhansky and Pavlovsky's experiment, would lead every F1 fruit fly to have the same genetic makeup and consequently, supposing a perfectly homogenous environment (in a deterministic setup), the same viability and fertility. In such conditions, no evolutionary change in terms of frequency of alleles would, in principle, be observed at the F2 generation. In practice, noise of different origins would lead to some change, but that does not invalidate the conceptual point I am making.

From an allelic point of view, at the F1 generation, all the token alleles of one type of allele would be systematically associated with one genetic background, while the token alleles of the other type with another genetic background. The variation between the two genetic backgrounds would only be due to the systematic presence of the other allele at the focal locus in the homologous chromosome of each organism since all fruit flies at that generation would be heterozygous (*AB*) at that locus. Thus, the only difference between allele *A* and *B* from the point of view of *A* would be that their allelic environment is systematically *B* while the allelic environment of *B* is systematically *A*.

Yet, at the generation F2 (which is the point at which Dobzhansky and Pavlovsky's experiment starts) things would change. At that generation, each type of allele would be associated with the two possible genetic backgrounds (due to the allelic variability on the homologous chromosome at the focal locus) half of the time. This is because, assuming that each fly is able to reproduce panmictically in Mendelian proportions (i.e. no segregation distorters), the proportion of each genotype at the focal locus would be 25% *AA*, 50% *AB* and 25% *BB*. Thus, for each *A* with a genetic background *A* there would be a *B* with a genetic background *A* (and reciprocally). From there, the effect of each allele on their own reproductive output, in

[15] For a discussion on the notion of screening-off see Brandon (1990).

each genetic background, could thus be assessed (at the F3 generation) because the proportions of the time *A* and *B* would be associated with *A* and with *B* respectively would be the same.

Suppose now a homogeneous environment for the F2 fruit flies. Following Dobzhansky and Pavlovsky's explanation, because the genetic variability between the samples of 20 and 4000 flies would be the same (all fruit flies in the founding population—whether small or large—being identical), the only effect observed would be the effect of natural selection in the two cases, despite a difference in initial founding population sizes. In fact, we have two genetically identical populations except on the focal locus. This, by definition, prevents the founder effect from occurring. In fact, for there to be some founder effect to begin with would require some genetic variation at the background locus.

Note that in real populations there would certainly be *some* other forms of drift with significant effects on the evolutionary change in the populations starting with 20 individuals different from that produced by the founder effect. Other forms of drift might include any circumstance that lead to one allele having a higher or lower frequency that cannot be causally linked to its effect on its reproductive output than the other allele (reviewed in Millstein 2016, for an interpretation of drift in terms of environmental variation see Bourrat 2017). These circumstances should be distinguished from the founder effect on which Reisman and Forber focus in their argument (see the next section for more on this point).

The example I proposed, although purely theoretical, flies in the face of the argument that were individual-level properties (e.g. changing the genetic background of token alleles) manipulated instead of population-level ones (e.g. changing the population size), no information would be gained by this manipulation, contra what is claimed by Reisman and Forber (2005, p. 1121). Everything else being equal, not only would the population-level manipulation, in this case, yield no difference, but the relation between the founding population size and variability in the frequency of the focal allele would be lost. On the contrary, varying individual-level variables, such as the genetic background of alleles, in some ways, could yield the deviations from expectation observed by Dobzhansky and Pavlovsky. Crucially, this means that if one were to choose founding populations of 20 fruit flies with a sum of genetic variation higher than in founding populations of 4000 fruit flies, one would find, everything else being equal, greater deviations from the expected change, at the focal locus, in populations with large founding sizes than in populations with small ones. This is the opposite prediction to what one would expect following the population-cause view.

As I have argued above this is in line with Dobzhansky and Pavlovsky's reasoning. Replaced in a more general context, Dobzhansky and Spassky (1962) who used a similar experiment setting as Dobzhansky and Pavlovsky have further corroborated this reasoning experimentally. They founded five populations of 20 heterozygous individuals that all descended from the *one* couple of homozygous individuals (*AA* for the males and *BB* for the females or vice versa), while five other populations of twenty heterozygous individual were founded from homozygous parents of different geographical origins. Dobzhansky and Spassky found that the evolutionary change in the populations that all descended from a single couple of parents

exhibited much less variation than the evolutionary change observed when the population descended from parents of multiple origins. The explanation given by Dobzhansky and Spassky is that in the former case there is much less genetic variation in the background and lead to a lower variation in outcome independently of the population size. This hypothesis has been supported by other studies reviewed in Dobzhansky (1970, pp. 249–257).

Going back to the criterion of invariance with respect to the founder effect, the explanation based on differences in genetic background (individual-level differences), in light of my thought experiment, is thus more invariant under intervention than an explanation based on founding population sizes. An individual-level manipulation that would lead the F2 generation to be more similar in their genetic background would lead to a decrease in drift *no matter what the founding population size is*, while an increase in founding population size would have no effect on allele frequencies if the population is clonal in the genetic background of the two focal alleles. It follows that for the same explanandum, namely explaining changes in allele frequencies, if one accepts both the manipulation condition and the invariance criterion, the individual-level explanation represents a better causal explanation for the case proposed by Reisman and Forber.

5 Population Level Causes in Indeterministic and Frequency-Dependent Cases

So far, I have assumed that biological processes are deterministic. While a number of philosophers believe that indeterminism is eliminable from evolutionary theory (see Rosenberg 2001, Sober 2010; Weber 2001 for discussions), something must be said about possible cases of drift resulting from indeterminism. Does the explanation in individual-level terms I provided in the previous section hold when differences in reproductive output are due to indeterministic processes?[16]

Take again the example starting with the clonal population of fruit flies in generation F1—homozygous at all the loci, except at the focal locus where they are *AB*. Suppose that the females of each type produce non-deterministically a given number of offspring (leading to a probability distribution for the number of offspring produced). Starting with populations of 20 flies (10 females), *everything else being equal*, deviations from expectations would be higher than when the initial population size is of 4000 flies (2000 females). Contrary to the deterministic case, manipulating initial population sizes would, in this case, have an effect on evolutionary change. This effect would decrease as the size of populations increases. The reason it would is because of a direct consequence of the ontologically probabilistic nature of processes leading to reproduction.

In such cases, because the manipulation condition could be applied between the population-level variable ‘initial population size’ and ‘evolutionary change’ *and* that

[16] We could, for instance, imagine that some quantum processes percolate up in a biological process and have consequences on reproductive outputs, as done by Glymour (2001).

no manipulation of individual-level variables would bring about the same outcome more invariably, it seems reasonable to consider drift as a population-level cause.

A similar reasoning can be applied with natural selection. If the differences in properties of individuals leading to differences in reproduction (or growth) are fundamentally probabilistic, given one event of reproduction at the individual level, this event could not be attributed to the effect of natural selection or drift, and natural selection can only be characterized at the population level. In this sense the difference between natural selection and drift is inscrutable (see Huneman 2014 for more on the notion of 'inscrutability' in this context). Although I do not demonstrate it here, one can remark however that in a deterministic setup in which there would be no *deterministic* drift (e.g., no founder effect following my interpretation in Sect. 4) explanations for why one variant invades the population in a case of *directional* selection would also be more invariantly explained by individual-level variable properties than population level ones. See for instance Godfrey-Smith 2009's recent attempt to provide such an explanation in terms of intrinsic properties' effects on reproductive output (see Bourrat 2015, 2017 for an update of this view).[17]

Finally, before closing, I should mention cases of *frequency-dependent* selection. Sarkar (2008), when discussing the unit of selection issue in different cases of frequency-dependent selection convincingly shows that in some simple cases of frequency-dependent selection such as the hawk-dove evolutionary games-theory model, fitness can be considered as an individual-level (allelic) property because it is definable at that level. In other cases, involving for instance heterosis or more complex evolutionary game-theory cases, it should rather be regarded as a property at a higher level (e.g., the genotypic level) because it cannot be defined solely at the lower level. This point can be extended recursively at any level of organization. If the conclusions reached by Sarkar are correct, it seems thus that for frequency-dependent cases (which are special cases of context-dependence), whether natural selection is an individual- or population-level property will depend on the details of the case. More particularly it will depend, for a given case, whether the evolutionary dynamics can be computed without references to higher-level variables.

This point leads us to Millstein's (2006) position. Like Reisman and Forber, Millstein explicitly considers natural selection and drift to be population level causes that do not reduce to individual level causes. To fuel this view, she proposes the case of a model with some experimental results taken from Kerr et al. (2002) of what they call a 'non-transitive' community of bacteria (*E. Coli*) in which there are three types: A, B and C. In the absence of C, A has a higher reproductive output than B. In the absence of A, B has a higher reproductive output than C. But in the absence of B, C has a lower reproductive output than A. This non-transitive case is analogous to the "rock-paper-scissors" game. This type of dependencies between types is a complex game-theory case, which, following Sarkar, cannot be defined purely at the individual level, since predicting the evolutionary dynamics in this case involves considering population-level properties.

[17] Needless to say, both deterministic and indeterministic natural selection/drift can be at play in a single population.

Thus, it is warranted, in such case and similar ones, to consider natural selection and drift as population-level causes of evolutionary change.

Millstein rightly notes that frequency dependence is considered in the literature as a common phenomenon. However, I do not believe that the conclusion that natural selection and drift are population-level causes *as opposed to* individual-level causes follows. Rather, I propose that depending on the case, the question being asked and the time-scale of the explanation, in brief, depending on the explanandum, natural selection and drift can be regarded as more or less population-level or individual-level causes. Holding a stronger position, namely that natural selection and drift are *necessarily* population-level causes, is problematic and leads to some slippery slope arguments. One might argue for instance that since any causal explanation (whether it concerns evolutionary change or not) is, to some extent, context-specific, then causes are always necessarily population-level causes. Such a position would make the defense of an individual-level cause view impossible. I believe that the protagonists of the debate all agree that in principle the two alternatives are viable ones and that whether a cause is an individual or population-level one refers to the locus of manipulability.

6 Conclusion

I have argued that Reisman and Forber's claims that natural selection and drift are population-level causes of evolutionary change because they (1) can be manipulated at the population level, (2) cannot be reduced to individual-level explanations, and (3) convey more information than individual-level explanations, are inadequate in the deterministic and frequency-independent cases they presented. In fact, starting from a similar scenario as the experiment they use, while assuming causal determinism, I showed that individual-level variables could be manipulated to produce the same effect as when manipulating population-level variables. I then showed how these variables can be related to the population-level variables used by population geneticists and explained why manipulating these population-level variables in deterministic set-ups is ultimately equivalent to manipulating individual-level variables. Later, I argued that if one uses Woodward's criterion of invariance, a causal explanation of the founder effect in terms of individual-level variables is more invariant under intervention than in terms of population-level variables and thus that, on that ground, the individual-cause is superior to the population-cause view. Finally, I conceded that natural selection and drift can be conceived of as population-level causes when they depend on indeterministic events and in some cases of frequency-dependent selection.

Acknowledgements I am thankful to Paul Griffiths, Frans Jacobs, Charles Pence, Arnaud Pocheville, Joeri Witteveen and three anonymous reviewers for their comments on earlier versions of the manuscript. This research was supported under Australian Research Council's Discovery Projects funding scheme (project number DP150102875) and a Macquarie University Research Fellowship.

References

Ariew A, Rice C, Rohwer Y (2015) Autonomous-statistical explanations and natural selection. Br J Philos Sci 66(3):635–658

Beatty J (1984) Chance and natural selection. Philos Sci 51(2):183–211

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

Bourrat P (2015) Distinguishing natural selection from other evolutionary processes in the evolution of altruism. Biol Theory 10(4):311–321

Bourrat P (2017) Explaining drift from a deterministic setting. Biol Theory 12(1):27–38

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

Dobzhansky T (1970) Genetics of the evolutionary process. Columbia University Press, New York

Dobzhansky T, Pavlovsky O (1957) An experimental study of interaction between genetic drift and natural selection. Evolution 11:311–319

Dobzhansky T, Spassky NP (1962) Genetic drift and natural selection in experimental populations of *Drosophila pseudoobscura*. Proc Natl Acad Sci 48(2):148–156

Fisher RA (1930) The genetical theory of natural selection. Clarendon Press, Oxford

Forber P, Reisman K (2007) Can there be stochastic evolutionary causes? Philos Sci 74(5):616–627

Futuyma DJ (2005) Evolution. Sinauer Associates Inc, Sunerland, MA

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

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

Haig D (2012) The strategic gene. Biol Philos 27(4):461–479

Hamilton M (2009) Population genetics. Wiley, Oxford

Hitchcock C, Velasco J (2014) Evolutionary and Newtonian forces. Ergo Open Access J Philos 1

Huneman P (2014) Inscrutability and the opacity of natural selection and random genetic drift: distinguishing the epistemic and metaphysical aspects. Erkenntnis 80(3):491–518

Kerr B, Riley MA, Feldman MW, Bohannan BJ (2002) Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors. Nature 418(6894):171–174

Lu Q, Bourrat P (2017) The evolutionary gene and the extended evolutionary synthesis. Br J Philos Sci. https://doi.org/10.1093/bjps/axw035

Matthen M, Ariew A (2002) Two ways of thinking about fitness and natural selection. J Philos 99(2):55–83

Matthen M, Ariew A (2009) Selection and causation. Philos Sci 76(2):201–224

Millstein RL (2006) Natural selection as a population-level causal process. Br J Philos Sci 57(4):627–653

Millstein RL (2016) Genetic drift. In: Zalta EN (ed) The stanford encyclopedia of philosophy (Fall 2016)

Northcott R (2010) Walsh on causes and evolution. Philos Sci 77(3):457–467

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

Okasha S (2008) Fisher's fundamental theorem of natural selection—a philosophical analysis. Br J Philos Sci 59(3):319–351

Otsuka J (2016) A critical review of the statisticalist debate. Biol Philos 31(4):459–482

Otsuka J, Turner T, Allen C, Lloyd EA (2011) Why the causal view of fitness survives. Philos Sci 78(2):209–224

Pearl J (2009) Causality: models, reasoning, and inference, 2nd edn. Cambridge University Press, New York

Plutynski A (2007) Drift: a historical and conceptual overview. Biol Theory 2:156–167

Pocheville A, Griffiths PE, Stotz K (2017) Comparing causes—an information-theoretic approach to specificity, proportionality and stability. In: Leitgeb H, Niiniluoto I, Sober E, Seppälä P (eds) Proceedings of the 15th congress of logic, methodology and philosophy of science. College Publications, London, pp 250–275

Reisman K, Forber P (2005) Manipulation and the causes of evolution. Philos Sci 72(5):1113–1123

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

Rosenberg A, Bouchard F (2005) Matthen and Ariew's obituary for fitness: reports of its death have been greatly exaggerated. Biol Philos 20(2):343–353

Sarkar S (2008) A note on frequency dependence and the levels/units of selection. Biol Philos 23(2):217–228
Shapiro L, Sober E (2007) Epiphenomenalism—the do's and the don'ts. In: Wolters G, Machamer P (eds) Thinking about causes: From Greek philosophy to modern physics. University of Pittsburgh Press, Pittsburgh, pp 235–264
Sober E (2010) Evolutionary theory and the reality of macro-probabilities. In: Eells E, Fetzer JH (eds) The place of probability in science. Springer, Dordrecht, pp 133–161
Stephens C (2004) Selection, drift, and the "Forces" of evolution. Philos Sci 71(4):550–570
Sterelny K, Kitcher P (1988) The return of the gene. J Philos 85(7):339–361
Walsh DM (2000) Chasing shadows: natural selection and adaptation. Stud Hist Philos Sci Part C Stud Hist Philos Biol Biomed Sci 31(1):135–153
Walsh DM (2007) The pomp of superfluous causes: the interpretation of evolutionary theory. Philos Sci 74(3):281–303
Walsh DM (2010) Not a sure thing: fitness, probability, and causation. Philos Sci 77(2):147–171
Walsh DM, Lewens T, Ariew A (2002) The trials of life: natural selection and random drift. Philos Sci 69(3):429–446
Walsh DM, Ariew A, Matthen M (2017) Four pillars of statisticalism. Philos Theory Biol 9(20170609)
Weber M (2001) Determinism, realism, and probability in evolutionary theory. Philos Sci 68(3):S213–S224
Woodward J (2000) Explanation and invariance in the special sciences. Br J Philos Sci 51(2):197–254
Woodward J (2003) Making things happen: a theory of causal explanation. Oxford University Press, New York
Woodward J (2010) Causation in biology: stability, specificity, and the choice of levels of explanation. Biol Philos 25(3):287–318
Woodward J (2013) Causation and manipulability. Stanford encyclopedia of philosophy. http://plato.stanford.edu/archives/win2013/entries/causation-mani/