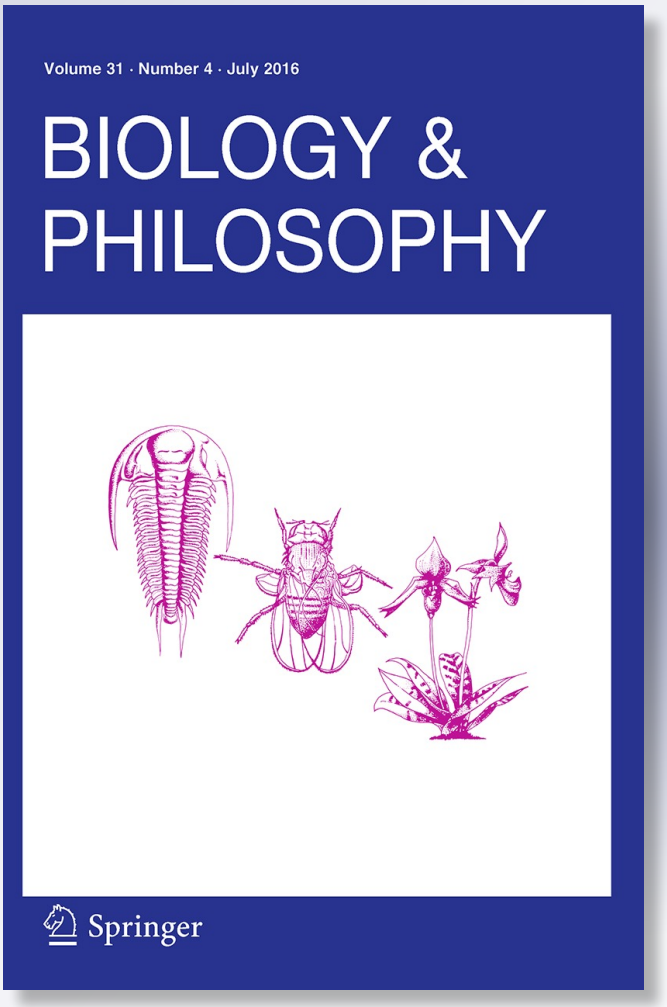
A critical review of the statisticalist debate

Jun Otsuka

Biology & Philosophy

ISSN 0169-3867
Volume 31
Number 4

Biol Philos (2016) 31:459-482
DOI 10.1007/s10539-016-9528-0



Springer

Biol Philos (2016) 31:459–482
DOI 10.1007/s10539-016-9528-0

AREA REVIEW

A critical review of the statisticalist debate

Jun Otsuka[1]

Received: 27 October 2015 / Accepted: 5 May 2016 / Published online: 24 May 2016

Abstract Over the past decade philosophers of biology have discussed whether evolutionary theory is a causal theory or a phenomenological study of evolution based solely on the statistical features of a population. This article reviews this controversy from three aspects, respectively concerning the assumptions, applications, and explanations of evolutionary theory, with a view to arriving at a definite conclusion in each contention. In so doing I also argue that an implicit methodological assumption shared by both sides of the debate, namely the overconfidence in conceptual analysis as a tool to understand the scientific theory, is the real culprit that has both generated the problem and precluded its solution for such a long time.

Keywords Statisticalism · Evolutionary genetics · Causal models · Selection · Drift · Fitness

Introduction

Understanding the nature of evolutionary theory is one of the central goals of the philosophy of biology. The traditional meta-scientific account describes evolutionary theory as a causal theory that explains changes of a population based on various causal factors including biotic as well as abiotic conditions. However, from the beginning of this century this view has come under criticism by a group of philosophers known as *statisticalists* (Matthen and Ariew 2002; Walsh et al. 2002). The statisticalists construe evolution as a purely statistical phenomenon, and argue that population genetics, the mathematical core of modern evolutionary theory, studies not causes of evolution but rather abstract relationships that exist among various statistics estimated by census. The statisticalists' claim has provoked a

✉ Jun Otsuka
jotsuka@lit.kobe-u.ac.jp

[1] Philosophy Department, Kobe University, 1-1 Rokkodai-machi, Nada, Kobe, Japan

number of critical responses by advocates of the traditional account, often called *causalists*. The purpose of this article is to critically review some major issues discussed in the debate, with a view to arriving at a definite conclusion in each contention.

During this decade of sometimes heated exchanges, the statisticalist-causalist debate has ramified to cover various topics including, to name a few, the interpretation of fitness, the conceptual distinction between selection and drift, the possibility of the population-level causation, and metaphorical representations of evolutionary theory. This review does not aim to cover all these contentions or relevant articles. In particular, I do not address the question as to whether evolutionary theory should be understood as a theory of *force* that compares the major factors of evolution such as selection and drift to Newtonian forces (Sober 1984). Although a criticism of this metaphor was among one of the earliest statisticalist agendas to put forward a non-causal view of evolutionary theory, these two issues are logically independent and thus can be examined separately (Stephens 2004; Lewens 2004). The primary focus of this review is the problem of the causal nature of evolutionary theory, so I address the above listed issues only in relation to this one main focus.

In the debate causalists have held, and statisticalists denied, that evolutionary theory is a causal theory. What, however, does it mean for a scientific theory to be (non) causal? This is already an important contention that lacks a clear consensus, but one may nevertheless discern in the discussion three distinct senses in which evolutionary theory, or any scientific theory in general, is claimed to be causal or non-causal. First, a theory is said to be causal (or non-causal) if it involves explicit (no) causal *assumptions*. According to statisticalists, the core principles of evolutionary theory, or at least of its formalized version after the Modern Synthesis, are mathematical theorems that hold regardless of causal details of populations. In response, causalists maintain the derivation of these principles does require some form or another of causal assumption. As the nature of a theory is largely determined by its premises, analyses and interpretations of the assumptions of evolutionary principles naturally go on to form the first contention in the statisticalist-causalist dispute. The second criterion of a causal theory related to but nevertheless distinct from the first is whether its *empirical application* requires any causal information about the target phenomena. What do we need to know about a population in order to predict its evolutionary trajectory? To be sure, we need some key measures such as fitness, but do we also need to know its causal basis? If so, it will make a strong case for the causal interpretation of evolutionary theory, and vice versa. Finally, a theory may be considered causal if its *explanations* are causal, that is, if they invoke causal relationships or concepts in an essential way. In the causalist picture, the goal of evolutionary theory is to identify causes of population changes, such as fitness variation. This is opposed by statisticalists who deny fitness to be a cause of evolution, and construe explanations in population genetics as subsumptions of target populations under some mathematical theorem.

Obviously these three criteria are not logically independent from each other: any causal explanation, for one, presupposes some knowledge about the alleged

cause, which would serve as an important theoretical assumption. Clarifying their logical connection in the debate is one of the goals of this review. Nonetheless they represent distinct strategies to argue for either the causal or non-causal interpretation of evolutionary theory, and for this reason the following review is structured according to these three criteria. We will first examine the dispute over the theoretical assumptions of evolutionary theory, followed by the contentions on empirical applications and the epistemic status of evolutionary explanations.

By setting this agenda, I by no means pretend for this review to be a neutral "view from nowhere." To the contrary, much of the following analysis is based upon the causal graph approach to evolutionary models that was developed in Otsuka (2016). The choice of this particular framework reflects my conviction that the problem and controversy at hand are best elucidated when viewed as one concerning the theoretical assumptions and constructions of evolutionary models. This, however, is not necessarily the way the question has been framed by the participants of the debate: rather it has most often been construed as a meta-scientific question over the correct *interpretation* of evolutionary theory or concepts. Such a methodological framing of the issue by itself reveals something about the debate, and in particular is not unrelated to the reason for its protracted struggle. This point will be elaborated in the last section, to draw a meta-philosophical diagnosis of the debate.

Round 1: The theoretical basis of evolutionary theory

Is evolution a mathematical necessity?

Modern evolutionary biology, like many other mature sciences, is highly mathematized. Ever since the Modern Synthesis, mathematical population genetics has served as the theoretical backbone and unifying principle for the development of evolutionary thinking in the 20th century. According to statisticalists, however, population genetics is not only mathematical—it is *a mathematics*. That is, not only does it deal with complex mathematical formulae (after all physics is full of sophisticated mathematics), but rather its principal equations describing evolutionary changes are all mathematical theorems, whose derivation requires nothing more than assumptions or axioms of, say, probability theory. Thus Matthen and Ariew (2009, p. 211) assert: "When there are heritable differences in traits leading to differential reproduction rates, the probability of the fitter types increasing in frequency is greater than that of the less-fit types increasing. *This is simply a mathematical truth*"(emphasis added). From this observation they conclude that "Mathematical population genetics is, in large measure, an application of probability/frequency theory."

What are these theorems that are said to govern evolutionary changes? One example featured by Matthen and Ariew (2002) is *Li's theorem* which gives the

change in the mean growth rate in terms of the variance of relative growth rates, or $\Delta\bar{W} = \mathrm{Var(W)}/\bar{\mathrm{W}}$, where the growth rate W of each type (e.g., allele) in the population measures the increase or decrease of its share in the next generation. Another example is the *Price equation* $\Delta\bar{Z} = \mathrm{Cov(Z, W)}/\bar{\mathrm{W}}$, which expresses the change in the phenotypic mean $\bar{Z}$ in terms of the covariance of the phenotype Z and the fitness (i.e., the number of offspring) W divided by the mean fitness $\bar{W}$. It is well known that the Price equation is an algebraic truth that holds just in virtue of the axioms of probability theory and the definitions of the mean, covariance, and the variables used therein.[1] The same is true of Li's theorem, which is a special case of the Price equation obtained by substituting phenotype Z in the Price equation for fitness or growth rate W and noting the covariance of a variable with itself is its variance.

If the general principles of evolution are a priori truths, predicting evolutionary changes is just a matter of algebraic calculation that requires no causal or even empirical assumption. On this ground statisticalists conclude "selection is mathematical in nature, and independent of the particular causal laws that produce growth" (Matthen and Ariew 2002, p. 74). This is not to deny that a selective episode consists of a host of causal interactions that culminate in individual births and deaths. The claim is rather that mathematical population genetics abstracts away all these causal substrates and studies selection as a purely mathematical relationship that can be described with a priori theorems. "Li's theorem tells us nothing about causes of growth: it is a general truth about growth regardless of how it is caused" (*ibid.*). Assumptions regarding causal details play no substantive role in establishing this "general truth" and thus in predicting evolutionary changes—hence follows the statisticalist doctrine that evolutionary theory is not a causal but purely statistical theory.

Causalist responses

Critics of the statisticalism have challenged this purely mathematical characterization of population genetics in two ways. The first line of response is to assert that statisticalists are looking at the wrong place to read off a causal implication of the theory. Millstein et al. (2009) criticize statisticalists for deriving their conclusion from the mere fact that selection is expressed by some mathematical formula. The fact that something can be represented with an a priori equation does not prove its non-causal nature, they argue, because it is not an equation itself but an *interpretation* that gives a causal content. Take their example of a binomial equation $(p+q)^2 = p^2 + 2pq + q^2 = 1$. This same equation can be thought of as representing either genotype frequencies at Hardy-Weinberg equilibrium, or the area of a unit square divided into four rectangles (one square with the size p^2, another with q^2, and two rectangles with pq). But the equation itself is silent as to which of these representations is correct: the representational content of the equation, and thus whether it represents a causal or physical process at all, is

[1] See e.g., Okasha (2006) for a derivation of the Price equation.

determined by its pragmatic context or the intention of the user who applies this equation to a particular problem. If so, that evolutionary equations turn out to be mathematical necessities would have no implication for the (non) causal nature of evolutionary theory—it is rather how they are used that counts. On this ground Millstein and colleagues argue that equations of population genetics, or more specifically the binomial representation of drift, have full causal meaning as they have been used by geneticists, most notably Fisher and Wright, to represent a class of causal processes which they call indiscriminate sampling process.

It is certainly right that the presence or absence of a mathematical expression alone does not determine the causal or non-causal nature of a given relationship, but why and how can an interpretation help in this regard? An interpretation maps a theory—a set of linguistic entities such as mathematical equations—to a particular domain like a set of squares or sexually reproducing populations. The premise of Millstein et al. (2009) is that this connection "infuses" the theory with the empirical or causal contents of the target domain. As a consequence, the contention on the nature of evolutionary *theory* is reduced to a metaphysical inquiry of evolutionary *phenomena*. It is in this context that Millstein distinguishes *discriminate and indiscriminate sampling processes*, as two types of causal processes that are affected or not affected by phenotypic differences between organisms (Millstein 2002, 2005, 2006). This is introduced as an ontological distinction holding prior to any mathematical formulation of evolutionary theory, and it is by being mapped onto these distinct processes that evolutionary concepts such as selection and drift acquire causal meanings.

I will postpone the examination of this ontological distinction until a later section, but even if we grant this distinction Millstein and colleagues are only half way through. Let's grant, for the sake of argument, that there are two distinct processes in nature. An important question, however, is: how can we know that these processes are the referents of the mathematical conceptions of selection and drift? To establish this semantic connection these processes must be shown to actually generate the evolutionary behaviors in question as quantitatively characterized by population genetics. On this regard Matthen (2010) asks: supposedly indiscriminate sampling is acting in an infinite as well as finite population, but then why does drift manifest only in the latter? Or in general, why does its action depend on the population size at all? To answer these questions one needs to "embed" the alleged processes within population genetics, identifying their place and role in the mathematical equations of evolutionary changes. Until this is complete one cannot conclude these processes to be the real world referents of selection and drift as conceptualized in evolutionary theory.

Moreover, Millstein et al.'s approach will not convince those statisticalists who think the issue in contention is ultimately of epistemological rather than ontological nature. Ariew and Ernst (2009) and Ariew et al. (2015), for example, detach the modern genetical theory from Darwin's original theory of natural selection, limiting the target of their non-causal claim only to the former while admitting the Darwinian theory addresses causes of evolution. They do not deny, therefore, evolutionary phenomena (a class of phenomena studied by theories of evolution) consist of causal processes: what they deny is that population genetics deals with

these causal relationships. Such a position is immune to Millstein et al.'s criticism, for Ariew and his colleagues can acknowledge the causal basis of evolutionary phenomena while denying the causal nature of population genetics as a theory, which they claim to study evolutionary phenomena after abstracting away all these causal contents.

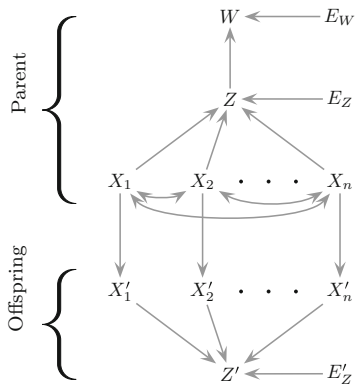
The above discussions suggest that to fully resolve the dispute one cannot avoid analyzing the mathematical structure of evolutionary theory. In this regard Millstein and colleagues may concede too much by accepting or at least not questioning the statisticalists' premise that the principles of population genetics are of purely mathematical nature. The second line of response challenges this premise. According to Rosenberg and Bouchard (2005), it is a mistake to think that the foundation of evolutionary theory is provided by mathematical formulae such as Li's theorem or Fisher's fundamental theorem of natural selection (FTNS). More fundamental than these equations is the following *principle of natural selection*:

> PNS (x) (y) (E) [If x and y are competing populations and x is fitter than y in environment E at generation n, then probably, (x's size is larger than y in E at some generation n' later than n)].

Rosenberg and Bouchard then claim that (i) the PNS is a causal principle, for it compares the *ecological fitness*, the causal capacity of individual organisms to survive and reproduce; and that (ii) the abstract formulae of evolution such as Li's theorem or Fisher's FTNS are all derived from this PNS. Taken together, they conclude that mathematical equations of population genetics, despite their abstract and purely statistical appearance, are in fact based on a causal principle.

I believe this approach to be on the right track, but remains incomplete for two reasons. First, to substantiate this claim Rosenberg and Bouchard must show that the FTNS or Li's theorem is actually derivable from their PNS, and despite their verbal promise that the derivation is "fairly direct and intuitive" it is far from obvious how a quantitative equation like the FTNS follows from a merely comparative principle like the PNS (we will return to this point later). Second, even if we put aside the feasibility of the derivation it is not clear in what sense the PNS is said to be causal. What exactly is the ecological fitness, and in what sense is it causal? Although we are told that the causal nature of the PNS stems from the concept of ecological fitness, "they do not tell us what this is," as Matthen and Ariew (2005, p. 359) complain.

The task for causalists, therefore, is to *actually* derive evolutionary equations from *explicitly causal* assumptions. Since most evolutionary equations are written in probabilistic forms, this requires one to connect two conceptually distinct realms, probability and causality. Although this problem has long been a source of bewilderment both for philosophers and statisticians, considerable progress has been made in the past few decades by the *causal graph theory*, which studies the formal relationships between causal structures expressed by directed graphs and probability distributions generated from them (Pearl 2000; Spirtes et al. 2000). Using this theoretical framework and Sewall Wright's *trek rule*, Otsuka (2016) identified causal models of evolving populations and derived the standard predictive equations of population and quantitative genetics. These causal assumptions include (1) a



$$W = \beta Z + E_W,$$

$$Z = \sum_i \alpha_i X_i + E_Z,$$

$$X'_i = X_i \quad \text{for } 1 \leq i \leq n,$$

$$Z' = \sum_j X'_j + E_{Z'}.$$

Fig. 1 A causal model underlying the breeder's equation. For the sake of simplicity the model here does not consider sexual reproduction or mutation. *Double-edged arrows* in the *graph* represent statistical dependence, or *linkage disequilibrium*, among parental genes. The structural equations on the *right* quantitatively specify each causal relationship in the *graph*

parent's alleles $(X_1, \ldots, X_n)$ affect its phenotype Z, which then contributes to the fitness W defined by the number of its offspring;[2] (2) the parental genes are passed down to offspring, which then affect the offspring's phenotype Z'; (3) environmental effects (E_W, E_Z, E'_Z) are independent; (4) all causal relationships are linear. The causal model thus defined (see Fig. 1) enables one to rewrite the breeder's equation $\Delta\bar{Z} = Sh^2$, which gives the between-generation response to selection, as a function of the causal parameters and the genetic variance such that:

$$\Delta\bar{Z} = \frac{1}{\bar{W}}\beta \sum_{i,j} \alpha_i \alpha_j \, \mathrm{Cov}(X_i, X_j)$$

where α and β are causal parameters of the structural equations (Fig. 1). More specifically, α_i measures the linear causal effect of the ith allele on the trait (the change in Z associated with a unit increase in X_i), while β is the causal effect of the trait on fitness (the change in W associated with a unit increase of Z).

The model can also be used to evaluate intervention effects on evolutionary responses, i.e., $P(\Delta\bar{Z}|do(Y = y))$ where $do(Y = y)$ denotes an intervention that sets the value of Y to y (Pearl 2000, see also section "Causes of evolution"). Although the causal model in Fig. 1 is the simplest case, Otsuka (2015) shows that this basic model can be extended to deal with more complex mechanisms such as epigenetic inheritance, maternal effects, and niche construction.

From a very general perspective, the goal of mathematical genetics can be viewed as identifying a function that returns the population change based on some information about the current population structure, such that

[2] Hence the model defines fitness as a measure of the *actual* number of offspring *causally related to* phenotype, contradicting the common philosophical wisdom according to which fitness is the *expected* offspring number that *supervenes* on phenotype and the environment. This discrepancy will be discussed in section "Much ado about fitness".

$$\text{evolutionary change} = f(\text{information about the current population}).$$

The statisticalism asserts the derivation of f does not require any more than the theory of probability or statistics. In the causal graph approach, however, the desired function is derived only from certain causal assumptions and the theory that explicitly handles them. That is, the predictive equations including the two-locus population genetics model, the breeder's equation, and the FTNS are all *theorems*, not of probability theory but of the causal graph theory which explicitly models causal relationships. Population genetics, therefore, is not "a mathematics" nor are its equations a priori truths; they are empirical propositions that hold only in virtue of certain causal assumptions. In this sense, the causal relationships are far from abstracted away but provide the very basis for the quantitative principles of evolutionary theory.

Evolutionary principles: explanatory or descriptive?

What, however, about the Price equation and Li's theorem? Aren't they mathematical truths? Yes they are, and as such their derivation does not require any causal assumption, as we have seen above. There is a catch, however—they are *not explanatory* at all. They may give a correct *description* of evolutionary changes, but not a *prediction* or *explanation* (e.g., Forber 2008). This becomes obvious if one takes a moment to look at, say, the Price equation $\Delta\bar{Z} = \mathrm{Cov}(Z, W)/\bar{W}$. Suppose you are to "predict" the change in the phenotypic mean, $\Delta\bar{Z}$, by calculating the right hand side. To do so you need to know the fitness W, i.e., the number of offspring, of each individual. If you know this, however, and assume perfect heritability, you also know the phenotypic distribution of the offspring generation, and thus the change between the two generations. Thus an application of the Price equation for the purpose of prediction would presuppose the very information you want to predict with it. In other words, it does not give a function f of the form above, for the right hand side of the Price equation involves information about the *next* generation. No causal assumption in, no prediction out—this is the reason why the Price equation or Li's theorem, being free from any causal assumption but hence devoid of predictive power, is seldom if ever used in empirical studies of adaptive evolution.[3]

For the very same reason the statisticalist accounts of evolutionary theory—which they call "hierarchical-realization scheme" (Matthen and Ariew 2002) or "statistically abstractive explanations" (Matthen 2009)—fail to capture the explanatory structure of the theory. According to these accounts, evolutionary explanations proceed by specifying the antecedents of mathematical theorems like Li's or Price's equation with concrete information about inheritance or selection of the population under study. Hence an instance of evolutionary explanation, say of the Galápagos finches in year 1970, is nothing but the most specific instantiation of the above a priori theorems. Statisticalists, however, have never shown a single example where biologists actually carry out such a subsumption-explanation. It does

[3] An empirical application of the Price equation can be found in Morrissey et al. (2012), but it is for the purpose of a post-hoc check of predicted adaptive responses, and not for predicting evolutionary response or detecting a selective pressure.

not come as a surprise, because that would be like turning lead into gold. What distinguishes the predictive models of population genetics from the mathematical identities such as Li's theorem is not levels of abstraction, but a set of additional causal assumptions—and without them no amount of concrete details turn the mathematical equations into explanations, at least in the sense used by most evolutionary biologists.

Therefore, a short answer to the statisticalist conundrum "how can population genetics be causal, if it is based on a priori principles like Li's theorem or Price's equation?" is "no, it is not based on a priori truths." Purely algebraic identities play no significant role in explaining evolution. It is rather causal assumptions about a population under study that enable the derivation of predictive equations of its evolutionary trajectories.

Round 2: Empirical applications of evolutionary theory

Although the causal reconstruction of the predictive equations discussed above highlights the importance of causal assumptions in population genetics, in a strict sense it does not disprove the statisticalist claim. It at most shows that those assumptions are *sufficient* to obtain the equations in population genetics, but not *necessary*. It is still logically possible, albeit very unlikely, that these or other causal assumptions turn out to be dispensable and any of these predictive equations proves to follow from purely mathematical axioms. Instead of engaging in such a direct proof, however, statisticalists have resorted to indirect arguments that focus on *empirical applications* of the equations, claiming evolutionary theory to be non-causal since its application to an actual population does not require any information about its causal features, but only statistical data. Indeed, if a theory makes some causal assumptions then it is natural to demand their confirmation prior to its application. Then by contraposition the dispensability of causal information would entail the non-causal nature of evolutionary theory.

Two arguments have been put forward in this line:

1. Causal analyses of the survival or reproductive capacity of organisms are at best comparative by nature and cannot yield the quantitative measure of fitness as used in population genetics (Matthen and Ariew 2002, 2009; Pigliucci and Kaplan 2006).
2. The causal features of a population are irrelevant in predicting its evolutionary change. Applications of evolutionary equations require only statistical information (Matthen and Ariew 2002; Ariew and Ernst 2009; Ariew et al. 2015).

In particular, the discussions have centered around the causal basis of fitness, which supposedly plays the central role in evolutionary explanations. In this context, the above two claims respectively assert the (1) impossibility and (2) dispensability of a causal analysis of fitness in applications of evolutionary theory. We will examine these claims in turn.

The formal and vernacular notions of fitness

As noted in the introduction the statisticalism emerged as an antithesis to the traditional meta-scientific account of evolutionary theory. The traditional account, in turn, was by a large part motivated by the notorious charge of tautology according to which evolutionary theory, summarized by "survival of the fittest," is a mere tautology that lacks any empirical content (Brandon 1978; Mills and Beatty 1979; Sober 1984, 1993). One of the best known strategies against the charge has been Sober (1984)'s distinction between *consequence laws* and *source laws*. The consequence laws are represented by equations of population genetics and calculate population changes with some quantitative parameters; while the source laws estimate these parameters based on behavioral or morphological features of organisms. Sober's solution to the tautology problem was that the alleged aprioriness of the consequence laws "does not hurt" the entire theory of evolution for its empiricalness is guaranteed by the process of applying the theory to an actual population via the source law. Likewise, the theory with a purely mathematical principle may be causal as a whole if its empirical application involves some sort or another of causal analysis, such as functional analysis of morphology.

Statisticalists needed to dismiss such a possibility in order to put forward their purely statistical interpretation of evolutionary theory. For this purpose Matthen and Ariew (2002) distinguished two concepts of fitness, vernacular and formal. The *vernacular* fitness is roughly what we have in mind when we say that one organism has an adaptive advantage over another in the Darwinian race of survival and reproduction. For any pair of organisms in a given environment, we can ask which is more adaptive or "fitter" based on their physical properties, say speed, body size, etc. According to Matthen and Ariew, this vernacular understanding of fitness is at best a comparative notion—e.g., one is faster, bigger, or stronger than another. In contrast, formal models of population genetics require a more elaborate measure of *formal fitness*, defined as "the *expected* rate of increase ⋯ of a gene, a trait, or an organism's representation in future generations"(p. 56, their emphasis). The per capita rate of increase is not just comparative, but comes in degree and is represented by a rational number. Now the problem they see in the Soberian solution is that there is a fundamental gap between a merely comparative order on the one hand and a quantitative measurement on the other: one can never arrive at the latter by comparison, but only by a direct census, they claim.

By the same token Matthen and Ariew (2005) reject Bouchard and Rosenberg (2004)'s use of the PNS to ground evolutionary principles on pairwise comparisons of organisms' capacity to solve a specific design problem posed by the environment. Such a capacity, or what Bouchard and Rosenberg call "ecological fitness," is nothing but the vernacular fitness in Matthen and Ariew's parlance, and for this reason they find it impossible to sustain quantitative formulae of evolution.

Be that as it may, are these two—comparative-vernacular and quantitative-formal—notions of fitness really inconsistent to each other? The contrary is suggested by measurement theory, a branch of applied mathematics that identifies operational criteria for assigning quantitative measures to a set of objects (e.g., Krantz et al. 1971). According to this theory, one of the most fundamental

requirements for objects to be measured with the ratio scale (which the "formal fitness" is) is that they allow pairwise comparison. This is intuitive if one recalls familiar measures, such as the kilogram system, are ultimately reduced to repeated pairwise comparisons by using, say, a balance. That the vernacular or ecological fitness is a comparative notion, therefore, is far from inconsistent but rather a necessary condition for a quantitative measurement of organisms' survival and/or reproductive performance.

In fact, under certain conditions repeated comparisons of reproductive success prove to be sufficient to give rise to the fitness measure as used in population genetics. Wagner (2010) devised such a pairwise competition test, where a pair of genotypes compete with each other with the "winner" being the one that increases its share against the other. Repeating the competition with different pairs of genotypes generates an order over the set of genotypes, upon which Wagner constructs a ratio scale measure of fitness and derives Wright's selection equation. This result substantiates Bouchard and Rosenberg (2004)'s idea to reduce the predictive measure of fitness to pairwise comparisons of reproductive or survival success, *pace* the statisticalist assertion that any such reduction is impossible.

The causal basis of fitness

Statisticalists, however, may be quick to respond as follows. Granted that Wagner's method allows us to construct a formal measure of fitness out of pairwise comparisons. What this method compares, however, are relative growth rates, not physical properties, of genotypes. Furthermore, since the growth rate of a genotype is estimated by census (i.e., by counting the number of its offspring), it is still a far cry from analyzing the formal fitness in terms of its causal basis.

The point is well-taken. Whether fitness is based on causal properties of organisms is an old question in the philosophy of biology, often debated under the heading of the *propensity interpretation* of fitness (Brandon 1978; Mills and Beatty 1979; Sober 1984, 2001, 2013; Rosenberg 1985; Brandon and Ramsey 2007; Ariew and Ernst 2009; Pence and Ramsey 2013). This labeling, however, may blur rather than reveal what the real issue is, for historically the "propensity interpretation" has been used by different authors to denote different theses, to name a few (i) that fitness is a *propensity or capacity* of an organism to survive and reproduce, rather than its actual performance; (ii) that fitness should be defined by the *statistical expectation*, rather than a sample moment; (iii) that fitness is *caused by organismal phenotype*; and (iv) that for any fitness function there is a *scalar value* that summarizes the direction of the adaptive response. Here we focus only on the third "interpretation" according to which fitness, as used in evolutionary theory, is a causal consequence of physical or behavioral properties of organisms. Statisticalists have challenged this thesis at two fronts: first, they deny an organism's fitness to be determined from its properties; second, it is argued that the fitness-phenotype relation need not be causal as long as there is a statistical association between them.

The first line of skepticism appeals to the context-dependency of fitness. In frequency-dependent selection the fitness of an individual organism depends on population-level parameters such as the population size or genetic/phenotypic

frequencies (Ariew and Lewontin 2004; Ariew and Ernst 2009). Gillespie (1974), for example, has shown that when a population consists of two genotypes reproducing at different variances, the evolutionary trajectory is affected by the population size (see also Frank and Slatkin 1990). Ariew and Ernst (2009) take this theoretical result to contradict the propensity interpretation of fitness in the sense defined above, for it shows a case where fitness cannot be uniquely determined from properties of an individual organism.

An obvious flaw in this argument is that proponents of the propensity interpretation do not need to assert the fitness of an organism to be determined *solely* from its own properties. All they need is that an individual property is *a*—not *the*—cause of fitness. To make an analogy, the premium of my car insurance is determined by, along with my own driving record, "population parameters" that summarize various conditions of hosts of drivers whom I haven't even met. Even still my driving record and habit affect my premium, and do so causally—it could have been less expensive should I have gotten less tickets, or used my car less frequently, and so on. Likewise, the fact that fitness depends on population parameters does not preclude an organismal character (either genetic or phenotypic) from being *a* cause of fitness.[4]

The second criticism of the propensity interpretation concerns the nature of the fitness-phenotype relationship. Even if fitness is a function of an organismal character, the functional relationship may not be causal, but just associational. Statisticalists in fact argue that it *need not be* causal, claiming adaptive evolution does not require anything more than differences in expected trait fitness, i.e., $E(W|z_i) \neq E(W|z_j)$ for different trait types $z_i \neq z_j$ (Walsh et al. 2002). This condition is of purely statistical nature and does not require type Z to be a cause of fitness W. For this reason statisticalists claim that modern evolutionary genetics does not concern causes, but only statistical properties of a population which can be estimated purely by census (Matthen and Ariew 2002; Ariew and Ernst 2009; Ariew et al. 2015).

This claim, however, is belied by actual practices of evolutionary ecology, one of the central concerns of which is to identify whether and how a phenotypic character under study *causally contributes* to the survival or reproductive success of organisms. Millstein's case studies (2006, 2008) show that field biologists take great pains to establish causal relationships between fitness on the one hand and phenotypic or genetic characters on the other in order to corroborate an adaptation hypothesis. Another classical example is Andersson's field study of sexual selection in which he found a positive effect of tail length of widowbirds on their mating chance (a surrogate measure of fitness) by experimentally manipulating the phenotype (Andersson 1982). Why do biologists insist on causality? The answer is because a mere phenotype-fitness correlation is not enough to induce an adaptive response (Glymour 2011; Otsuka 2016). Despite statisticalists, the essential condition for a trait to change its frequency in the subsequent generation in response to selection is not just a correlation or difference of expected trait fitness, but that the trait *causes* fitness. Two

[4] For this reason some statisticalists, e.g., Walsh (2007, p. 288), avoid taking this approach.

conditions, statistical and causal, come apart when there is a confounder, e.g., an environmental factor that affects both fitness and the phenotype. In such cases no evolutionary response occurs even if fitness correlates with the phenotype, and thus a prediction based just on differential expected trait fitness will be biased (Rausher 1992; Morrissey et al. 2010). Such biases can be eliminated only by confirming the observed fitness-phenotype correlation is fully accounted for by the causal effect of the trait, and for this reason information about causes is indispensable in empirical studies of evolution.

The apparent autonomy

In a recent defense of the statisticalism Ariew et al. (2015, pp. 647–648) claimed:

> in each case [of explanation in population genetics] the explanans is 'statistically autonomous', involving two general steps: assumptions that allow for the use of a statistical model and then deduction from that model . . . this deductive procedure is sufficient for explanation and no further appeal to causes is necessary.

That is, all we need to know to apply population genetics models are "the statistical properties of the population—for example, its mean and variance (p. 651)" but not the causal properties. This view is supposedly motivated by the fact that population genetics models are usually expressed in terms of statistical functions. However, since these equations are derived from certain causal assumptions (section "Round 1"), a violation of these causal assumptions may result in a wrong prediction, *even if one gets all the relevant statistics right.*

Any model is only as good as its assumption. For a successful application of an evolutionary model, researchers must take the following steps: (1) choose an appropriate model for a target population based on its biological and environmental conditions; (2) verify that the population satisfies the causal assumptions specified by the model (e.g., mating system, selective regime, etc.); (3) estimate the model parameters via statistical methods such as regression or analysis of variance. The statisticalist claim that explanations in population genetics are "statistically autonomous"—that they require only statistical information—stems from an exclusive focus on the last step. The apparent "autonomy" is illusory in two senses. First, these statistics are in fact estimates of the causal parameters (e.g., parameters in the structural equations). Second, the justification that such statistical functions correctly predict evolutionary changes can come only from the veracity of the underlying causal assumptions (steps 1 and 2). Hence far from being unnecessary, an "appeal to causes" is crucial in empirical applications of the mathematical models, and for this reason biologists take pains to identify the causal structure of a population in evaluating selection hypotheses or predicting future evolutionary trajectories (e.g., Wade and Kalisz 1990; Morrissey et al. 2012).

Round 3: Evolutionary explanations and interventions

Thus far we have discussed the statisticalist controversy from two perspectives, one regarding the theoretical structure and the other empirical applications of evolutionary theory. The debate has yet another face, which concerns the nature of evolutionary *explanations*—does the theory provide causal explanations of population changes?

To answer this question we must first ask when an explanation in general is considered causal. Traditionally philosophers have treated a scientific explanation as a relationship between two sets of propositions, *explanans* and *explanandum*. An explanation is called causal if the former identifies a cause of the phenomenon described by the latter (Sober 1984, ch. 5). Statisticalists have thus argued that the *explanans* of evolutionary changes refers only to statistical, but not causal, features of the population. This, as we have seen, was the gist of Ariew et al. (2015)'s claim that evolutionary explanations are "statistically autonomous."

In response, causalists have tried to show that evolutionary explanations indeed identify causes of evolutionary changes. Many of such arguments resort to the *interventionist account of causation* (Woodward 2003), according to which some variable X is a cause of another Y if there is a hypothetical intervention on X that changes the probability distribution of Y.[5] Thus the standard causalist strategy has been to point to a manipulation of selection, fitness, or drift that affects population frequencies. Sober and Shapiro argue that manipulating fitness or the variance thereof makes a difference in evolutionary response (Shapiro and Sober 2007; Sober 2013). Reisman and Forber submit a similar argument with respect to drift, arguing that an intervention on the population size affects the strength of the drift (Reisman and Forber 2005; Forber and Reisman 2007).

These claims did not go unchallenged. Statisticalists criticized such putative interventions as they do not satisfy some criterion or another of the interventionist account, and thus fail to establish the causalist conclusion. These challenges are examined in detail below.

Walsh's description independence thesis

In the Sober-Shapiro approach, the key contention is whether fitness can be a cause of adaptive evolution.[6] This has been put into question by a series of papers by Walsh (2007, 2010, 2015), who claims fitness fails to satisfy a necessary criterion of being a cause.

Walsh's argument is inspired by a well-known statistical puzzle called Simpson's paradox. Suppose two variables X and Y, and some partition of a population. Our intuition tells us if X and Y are positively correlated within every subpopulation, they must be so too in the overall population. This expectation is belied—the sign of

[5] The precise definition given by Woodward (2003) is more nuanced than this due primary to a possible violation of faithfulness, but these details can be ignored here.

[6] Note that this differs from the question regarding the causal basis or propensity interpretation of fitness as discussed above (section "The causal basis of fitness"), which asks whether phenotype can be a cause of fitness.

correlation can flip between sub- and whole-population. Such an apparent paradox has long been recognized by statisticians as well as philosophers, but Walsh gives a twist to this puzzle by maintaining that the phenomenon is peculiar only to merely statistical associations. That is, he claims in cases where X causes Y Simpson's reversal cannot happen: if, for example, X positively contributes to Y in each subpopulation, it must do so too in the whole population. Walsh (2007) calls this "description independence" of causal relationships, and seeks its justification in Judea Pearl (2000)'s *Sure Thing Principle* (Walsh 2010).[7]

The second step of Walsh's argument is to show that under a certain circumstance fitness does not satisfy this context independence. The case in question is again Gillespie (1974)'s model discussed earlier: if two types, say A and B, reproduce at different variances their long-term growth rates depend not only on the individual performance of each type but also on the population size. The moral Walsh draws from this is that if one describes the competition in small subpopulations A might be fitter than B, while in the whole population the opposite may hold. Fitness, hence concludes Walsh, is not description independent and thus cannot be a cause of evolution.

There are some confusions in Walsh's argument, most notably that the growth rate of a genotype in the Gillespie model is different from its fitness. This aside, there are fundamental errors in both of the two premises of Walsh's alleged *reductio*, namely that (1) causal relations must be description independent and that (2) Gillespie's model generates Simpson's reversal. With respect to (1), Northcott (2010) points out Walsh's description independence applies only to additive causes—in non-linear cases causal effects do depend on the background context, not only in magnitude but also in sign. Similarly, Otsuka et al. (2011) demonstrate Walsh's justification of his criterion by Pearl (2000)'s Sure Thing Principle stands on a misunderstanding of Pearl's theory of causality, and is unsound. Taken together we see the description independence is far from a necessary condition for causal relationships and cannot serve to disqualify the causal power of fitness.

Otsuka et al. (2011) also note (2) the alleged "fitness reversal" in Gillespie's model obtains only under an invalid assumption that one can set the population size in an arbitrary way, as if whether an organism belongs to the larger or smaller population is a matter of description. Such a supposition, however, not only incoheres with the construction of the Gillespie model which is derived for a predetermined population size, but also contradicts biologists' general wisdom that the (effective) population size is an objective feature of a population that must be estimated, but not be arbitrarily determined, by researchers. The last point cannot be emphasized enough, for the correct estimation of the size of an evolving population is one of the most important and challenging problems for modern population genetics since Fisher and Wright (e.g., Caballero 1994; Coyne et al. 1997; Wade and Goodnight 1998; Lynch 2007). If it were purely a "matter of description," these efforts for estimation and debates over *the true* population size would lose their entire meaning.

[7] Though not mentioned in his paper, a criterion similar to Walsh's was already proposed by Cartwright (1979) and criticized by Dupré (1984).

Some causalists have resisted this realist take on population size, pointing out that scientists reserve the right to choose (the size of) a population to be studied (Abrams 2013; Ramsey 2013). It is true, or even truism, that scientists can and *must* decide on which population they are going to investigate, and their decision surely reflects varieties of epistemic as well as pragmatic factors such as research interests, available resource, considerations on statistical power, etc. Abrams cites cases of selection study on human populations that pooled some subpopulations for the sake of a greater statistical power. Yet another research group may well prefer a smaller population due to sparse data or limited resources. Such decisions must be made, but that's not the end of the story. They must further be *justified* vis-a-vis their research goal, and such justifications can come only from nature. Pooling populations is allowed only when they are homogeneous (no mixture distribution), and extrapolations from a small population always risk overgeneralization—and all of that depends on the causal structure. Hence although it is scientists who demarcate the population to be investigated, whether their decision turns out to be correct is not up to them.

Before moving, let us note that the statisticalist supposition that a population can be demarcated in arbitrary ways is a logical consequence of their doctrine that evolutionary equations are purely mathematical truths. Indeed, nothing prohibits one from applying the Price equation to a gerrymandered population. Suppose a "population" consisting of all American citizens whose first name start with "T," all kangaroos living in Queensland, and my three goldfish. Count their descendants at some later time and the Price equation gives the exact change of any arbitrarily chosen phenotypic mean, say height (length). This is precisely because the Price equation, as an a priori mathematical theorem, applies to whatever set of objects as long as they satisfy certain measurement conditions. This is not true with *predictive* evolutionary equations such as the breeder's equation, whose derivation requires certain causal assumptions. To apply these equations, a population must be homogeneous with respect to the causal structure and consistent with the model assumptions. A causal structure is the unit of evolutionary theory which both affords and delimits the generalizability of evolutionary equations, and for this reason an evolving population cannot be demarcated willy-nilly.

Fitness-evolution relationship: causal or identical?

Recall under the interventionist account X counts as a cause of Y if there is an intervention on X that changes $P(Y)$. Based on this idea the causalists have suggested that interventions on fitness (Shapiro and Sober 2007; Sober 2013) or on the population size (Reisman and Forber 2005; Forber and Reisman 2007) affect adaptive evolution and drift, respectively. In order to establish the causal relation, however, the hypothetical interventions must satisfy an additional condition: namely that the intervened variable X and the supposed effect Y cannot be logically related. Manipulating a man's marital status would certainly change whether he is a bachelor or not, but it is not because they are causally related, but rather logically the same. According to Matthen and Ariew (2009), the same applies to the interventions proposed by the causalists. Although manipulating, say, the fitness

variance may affect evolutionary changes, it is just because they are the logically same thing—a variation in fitness *is* evolution.

The argument they develop to support this claim may be summarized into two points. The first is the now-familiar statisticalist doctrine that evolutionary equations relating fitness variation to evolution are mathematical truths—"natural selection is mathematically necessary" (Matthen and Ariew 2009, p. 211). As we have already seen, however, they are *not* mathematical truths, thus this line of reasoning may be dismissed. The second point concerns their peculiar definition of selection: "natural selection is evolution due to heritable variation in fitness" (Matthen and Ariew 2009, p. 204).[8] Defined in this way, of course selection logically implies adaptive evolution, but concluding the causal inertness of selection on this ground is just moving the goalposts. In fact, their "definition" of natural selection is a far cry from its common usage and contradicted by the very opening sentence of Fisher (1930): "Natural selection is not evolution."

Causes of evolution

Although the charge made by Matthen and Ariew (2009) may be dismissed as ungrounded, this does not automatically vindicate the causalism. To prove some variables to be causes of evolution, it must be shown that an intervention on those variables is well defined and effectively affects the evolutionary response. How can this be achieved? In the causal graph theory an intervention is represented as a manipulation of a causal model, and using the manipulated model the effect of the intervention can be evaluated in a straightforward manner (Pearl 2000; Spirtes et al. 2000). Hence the causal model underlying the breeder's equation (Fig. 1) may be used to examine if a variable of interest, such as fitness, causes population changes. Otsuka (2016) identifies two types of intervention affecting linear evolutionary changes. First, manipulating selective pressure β affects the rate and direction of evolutionary responses by regulating the contribution of the phenotype to fitness. Second, so-called "soft interventions" (interventions that leave other causal inputs intact) on fitness influence the rate of evolutionary changes. Suppose, for example, the skin thickness of some lizards contributes to their fitness by functioning as thermoregulation. Then raising (or decreasing) the environmental temperature will lead to a negative (positive) response in the mean skin thickness, with the rate of evolutionary change being proportional to the absolute value of the temperature change. On the other hand, culling a certain number of offspring of each individual regardless of its skin will not affect the direction of response, but will rather accelerate adaptive evolution of the skin thickness. These interventions, therefore, unambiguously identify causes of adaptive evolution.

Although Otsuka (2016) focuses exclusively on selection, drift can be shown to have a cause in a similar manner. In a linear selection model the strength of drift is measured by the variance of the average phenotypic change, $\mathrm{Var}(\Delta\bar{Z})$, where the upper bar denotes the sample mean in this context. For the sake of simplicity let us

[8] Matthen and Ariew attribute this definition to Sober (1984, pp. 21–22), but I couldn't locate it in the pages they point to.

focus only on the randomness in fitness, assuming perfect heritability (i.e., $\mathrm{Var}(E_Z) = \mathrm{Var}(E'_Z) = 0$ in Fig. 1). Then the expected strength of drift is obtained by calculating the variance of the sample covariance in the Price equation $\Delta\bar{Z} = \mathrm{Cov}(W, Z)/\bar{W}$, which yields:

$$\mathrm{Var}(\Delta\bar{Z}) = \frac{1}{N\bar{W}^2}\mathrm{Var}(Z)\mathrm{Var}(W) \quad (1)$$

(Rice 2004, pp. 183–185). This equation identifies three factors contributing to drift: population size *N*, phenotypic variance Var(Z), and fitness variance Var(W). Using the structural equation for fitness in Fig. 1, the last factor is unpacked into

$$\mathrm{Var}(W) = \beta^2\mathrm{Var}(Z) + \mathrm{Var}(E_W). \quad (2)$$

Combined with Eq. 1, this means one can regulate the strength of drift by manipulating the independent error term E_W.

The independent error term summarizes all causes of fitness that are independent of and act additively with respect to the phenotype in question. Examples along the line of the above hypothetical lizards might include predators' attack and mating chance, provided these factors do not interfere or correlate with the thermoregulation of the skin.[9] These additive factors of fitness correspond to what Millstein (2005, p. 171) calls *indiscriminate sampling process* "in which physical differences between organisms are causally irrelevant to differences in reproductive success," with two reservations. The first is that indiscriminateness of a sampling process is always relative to a particular feature. Though predators' attack may be indiscriminate with respect to skin thickness, it may not and need not be so as regard to other traits such as speed.[10] Secondly, the causal irrelevance is interpreted as a lack of G × E interaction. Thus that predators are indiscriminate sampling of the skin thickness means the trait does not affect or regulate the sampling process so that whether you have thick or thin skin does not raise or lower your chance of getting eaten. Provided these caveats, our conclusion here substantiates the claim of Millstein et al. (2009, see also section "Causalist responses") that the indiscriminate processes underlie drift. On the other hand, the selective pressures that regulate the fitness contribution of the focal phenotype correspond to a *discriminate sampling process* "in which physical differences between organisms are causally relevant to differences in reproductive success."

It has long been an issue whether selection and drift should be understood as mere "outcomes" or "processes" (e.g., Walsh et al. 2002; Stephens 2004; Matthen 2009, 2010; Millstein 2002, 2005; Brandon 2005). Proponents of the mere-outcome view hold selection and drift to be indistinguishable in terms of causal processes or forces, the distinction emerging only as a result of statistical abstraction. But if

[9] However, since interventions on these error terms count at the same time as soft-interventions on fitness, they may also affect the rate of adaptive response if they change the mean fitness. I thank Bruce Glymour for pointing this out.

[10] Speed in this case, however, must be causally disconnected (or *d-separated*; Pearl 2000) from the skin thickness for predators' attack to be an indiscriminate sampling of the latter. Otherwise there is an indirect *selection of* skin thickness (Sober 1984).

selection and drift appear them to be irreducible statistical facts, this is because they exclusively focus on the equations which only describe evolutionary outcomes. It is by uncovering the causal basis of these equations that we can find the causal processes governing evolutionary changes.

Individual versus population level causality

It should be noted that the causes of evolutionary changes as shown above all belong to *the level of individuals*, in the sense that these variables, including environmental factors, denote properties of individuals.[11] In contrast some causalists have argued that selection and drift should be understood as population-level causes that act on an entire population (Reisman and Forber 2005; Millstein 2006; Abrams 2007; Shapiro and Sober 2007). The attention to the population-level causation has been prompted as a response to the statisticalist syllogism of the following form:

(P1) Evolutionary theory deals with population properties.
(P2) Population properties are causally inert.
(C) Therefore, evolutionary theory is not causal.

Note the structural similarity of this argument to the *causal exclusion problem* in philosophy of mind, in which psychological explanations based on mental properties are claimed to be non-causal because mental properties are only epiphenomena of their physical or neural correlates.[12] The standard response to the causal exclusion problem has been to deny the epiphenomenalism and bestow some causal autonomy to the mental (e.g., Robb and Heil 2003). Likewise, the causalists have tried to disprove the ephiphenomenolist premise (P2) of the above statisticalist argument by asserting that selection (Millstein 2006; Abrams 2007), fitness (Shapiro and Sober 2007; Sober 2013), or population size (Reisman and Forber 2005) is a population-level cause of evolutionary changes. The contention then becomes how can these alleged population properties have autonomous causal power *irreducible to* individual births and deaths (Matthen and Ariew 2009; Matthen 2010), with a consequence that the whole statisticalist debate is conceived as if it is a variant of the causal exclusion problem—i.e., whether macro (population/mental) properties can have causal power beyond micro (individual/physical) properties.

That, however, is not the only possible response to the statisticalist syllogism: one could rather deny the first premise (P1) that evolutionary theory deals entirely with population properties (e.g., Bouchard and Rosenberg 2004). According to the causal graph approach featured in this review, evolutionary models are causal models of individual organisms that represent how each member of the population develops, survives, and reproduces in a given ecology and a genetic background. Hence unlike psychological theories, and contrary to the statisticalist and causalist presupposition alike, evolutionary theory does not abstract away, but rather explicitly deals with causal structures of individual organisms. Its causal nature

[11] Environmental factors in the causal graph represent these aspects of environment that are "experienced" by each individual, and are properties of individuals in this sense.

[12] I thank Patrick Forber for bringing my attention to this similarity.

derives from this fact, and does not depend on the contentious causal efficacy of macro-level properties, whatever they are. The apparent similarity to the causal exclusion problem, therefore, is a red herring that distorts the real issue of the statisticalist debate—it is a problem about the structure of a scientific theory, not a metaphysics of macro-level causation.[13]

Much ado about fitness

Now we have completed the examination of the statisticalist debate in relation to theoretical assumptions, empirical applications, and explanations of evolutionary theory. That, however, is not the way the issue has been framed by the disputants: rather it has been construed as a controversy over interpretations of key concepts in evolutionary theory, most notably that of *fitness*. According to the philosophical wisdom, fitness is a *propensity* of an organism that *supervenes* on its phenotype and environment (Brandon 1978; Mills and Beatty 1979; Sober 1984). As is well known, this view was initially devised as a response to the infamous charge of tautology, which claims the Darwinian theory, summarized as "survival of the fittest," is a tautology without empirical content. Isn't it circular to explain an organism's fate based on its fitness, if it is defined by the organism's actual survival or number of offspring? The standard strategy against this charge is to distinguish the potential capacity from actual performance, defining fitness as some latent propensity or disposition that supervenes on the phenotype and environment. The fitness thus defined, it is claimed, can explain the life and death of an organism without circularity, just like fragility of a glass, which supervenes on its physical makeup and the environmental conditions such as the medium-viscosity or temperature, may well explain its eventual shattering.

To the mind accustomed to this tradition, our definition of fitness W as the *actual* or *realized* number of offspring *caused* by phenotype (see footnote 2) might appear to fall prey to the old problem. This concern proves unfounded, not because our definition fares well with the problem, but rather because there exists no problem in the first place. As I see it, the whole tautology debate was prompted by the erroneous conception that fitness must explain the fate of an organism. It need not. For one, the explanatory target of evolutionary theory is not survival and/or reproduction of organisms, but *evolution* of a certain trait. And there is no tautology in saying that a trait will spread in the next generation because individuals having that trait better survived or had more offspring—indeed the proposition can be false if, among other things, the trait has zero heritability or the fitness-phenotype relationship is confounded (see section "The causal basis of fitness"). Hence I second Birch (2014) that the tautology problem is a pseudo-problem: evolutionary theory is empirical from the outset and does not rely on a particular interpretation of fitness to *make* it empirical. The illusion of tautology stems from the failure to identify the explanatory objective of the theory. In this regard, the defendants

[13] This does not mean, however, the causal exclusion problem does not exist in evolutionary biology. A similar problem, for example, emerges in the debate over levels of selection, i.e., whether group properties may have fitness contribution irreducible to individual properties (Okasha 2006). This is an open interesting question, but should be distinguished from the statisticalist debate.

as well as plaintiffs of the tautology debate have both committed to the same mistake made by the statisticalists, namely the confusion of descriptions and explanations (section "Evolutionary principles: explanatory or descriptive?"). Although "survival of fittest" may be an admissible (but vainly pedantic) *description* of selection, it falls a long way short of representing the kind of *explanation* provided by evolutionary theory. That the slogan might be tautological thus poses no threat to its epistemic agenda.

There is yet another deeper, structural similarity between these two debates, which reflects the oft-implicit assumption in the field that has only served to entangle the issues. As mentioned above, both debates have been framed by philosophers as the problem about the concept of fitness (or selection, drift, etc.), as if the empirical or causal nature of evolutionary theory hinges on its correct interpretation. Why does interpreting concepts have anything to do with the theoretical nature? One implicit rationale, I suspect, is the belief that evolutionary theory can be summarized into a few slogans like "heritable fitness variation leads to adaptive evolution" or "survival of the fittest." Taking them at face value, it was hoped that clarifying the key concepts in these explanantia would help us understand the epistemic nature of the theory. If fitness is what does the explaining of population changes, the nature of the explanation must be determined by its conceptual analysis—such reasoning led philosophers to the exegetical study of fitness to establish the empirical, causal, or statististical nature of evolutionary theory.

True, the common lore holds that evolutionary theory explains adaptive changes by fitness variation. Casting the theory into a few concise sentences like this, however, blurs the fact that most explanatory works in evolutionary science proceed by building models of target systems (Lloyd 1988). From this perspective, the straightforward way to identify the nature of evolutionary explanations is to analyze the construction process and assumptions of these models, rather than to interpret their verbal recapitulations. I thus believe evolutionary theory does provide causal explanations, but it is not because "it cites a cause of evolution," or even because some interventions on fitness affect evolutionary changes, but because it relies on models that deal explicitly with the causal structure of evolving populations. It is these theoretical contexts that determine the causal implication of the concepts like fitness, selection, or drift, and not the other way around. It is no wonder, therefore, that the significant amount of interpretative works which have accumulated in the past few decades have been to no avail in settling the debate. The failure, in my view, has a structural cause: it is the very paradigm of the philosophical practices that precluded the solution, and even created the problems themselves. Getting over the statisticalist and tautology controversies requires not just a new solution, but rather an alternative methodology for doing philosophy of biology.

Conclusion

Statisticalists hold that evolutionary theory is not a causal but purely statistical theory. The present review critically examined this claim from three perspectives, each concerning the assumptions, applications, and explanations of evolutionary theory. From any perspective the statisticalist doctrine cannot be maintained.

Contrary to the claim that evolutionary changes are mathematical necessities, deriving predictive equations in population genetics requires not just probability theory but also causal models and assumptions. An empirical application of any of these equations is thus contingent upon the causal features of the target population. This also means that evolution is explained from these causal features, with adaptive as well as non-adaptive evolution having corresponding causes in the sense of the interventionist account of causation.

In *Critique of the Pure Reason* Kant emphasized the importance of formulating a question in the right way, comparing attempts to answer ill-formed questions to "milking a billy-goat while the other holds a sieve underneath" (Kant 1998, p. 197, A58). I think a similar moral applies to the debate under review. How should we ask, if we are to investigate the causal nature of evolutionary theory? From the beginning the statisticalist controversy has been framed as a problem about *interpretations* of fitness, selection, or drift. Matthen and Ariew (2002) alleged an inconsistency between two interpretations of fitness—vernacular and formal—while Walsh et al. (2002) aimed to "distinguish dynamical and statistical interpretations of evolutionary theory." To these challenges causalists have responded with counter interpretations, such as the propensity view of fitness or the process view of selection/drift.

Implicit in this strategy is the meta-scientific belief that the nature of evolutionary theory will be elucidated by analyses of the key concepts appearing in its explanatory statements or summaries. One should not confuse, however, a summary with the theory. The popular principles or equations of evolution do not stand alone but are derivative of underlying models, and the concepts or parameters lose their meaning if detached from the theoretical context. To neglect this and ponder just about interpretations of linguistic expressions is like holding a sieve under a billy-goat, whereas what one should really examine is the goat, i.e., the model, itself! Once we turn our attention to the construction process of the models used in population genetics, it instantly becomes evident that they are far from a priori but based on causal, and thus empirical, assumptions. At the same time, the meaning and nature of the concepts like fitness, selection, and drift are determined unequivocally within these causal models.

The question about the causal nature of evolutionary theory is not about its interpretation, but rather about its models or the theory itself. Hence to adapt Patrick Suppes' famous slogan, the problem is properly addressed by a scientific, rather than meta-scientific, analysis. A philosophical exegetics unaccompanied by a serious analysis of the theory itself fails to establish a secure conclusion but leads only to an endless disputation. This, I think, is the lesson we should draw from the debate that has lasted for over a decade.

Acknowledgments I thank Patrick Forber, Charles Pence, and the members of Griesemer/Millstein philbio lab at UC Davis for their helpful comments on earlier versions of this paper.

References

Abrams M (2007) How do natural selection and random drift interact? Philos Sci 74(5):666–679

Abrams M (2013) Populations and pigeons: prosaic pluralism about evolutionary causes. Stud Hist Phil Biol Biomed Sci 44:294–301

Andersson M (1982) Female choice selects for extreme tail length in a widowbird. Nature 299:818–820
Ariew A, Ernst Z (2009) What fitness can't be. Erkenntnis 71:289–301
Ariew A, Lewontin R (2004) The confusions of fitness. Br J Philos Sci 55(2):347–363
Ariew A, Rice C, Rohwer Y (2015) Autonomous-statistical explanations and natural selection. Br J Philos Sci 66:635–658
Birch J (2014) Hamilton's rule and its discontents. Br J Philos Sci 65:381–411
Bouchard F, Rosenberg A (2004) Fitness, probability and the principles of natural selection. Br J Philos Sci 55:693–712
Brandon RN (1978) Adaptation and evolutionary theory. Stud Hist Phil Sci A 9(3):181–206
Brandon RN (2005) The difference between selection and drift: a reply to Millstein. Biol Philos 20(1):153–170
Brandon RN, Ramsey G (2007) What's wrong with the emergentist statistical interpretation of natural selection and random drift? In: Hull DL, Ruse M (eds) The cambridge companion to the philosophy of biology. Cambridge University Press, Cambridge, pp 66–84
Caballero A (1994) Developments in the prediction of effective population size. Heredity 73:657–679
Cartwright N (1979) Causal laws and effective strategies. Noûs 13(4):419–437
Coyne JA, Barton NH, Turelli M (1997) Perspective: a critique of Sewall Wright's shifting balance theory of evolution. Evolution 51(3):643–671
Dupré J (1984) Probabilistic causality emancipated. Midwest Stud Philos 9(1):169–175
Fisher RA (1930) The genetical theory of natural selection. Oxford University Press, Oxford
Forber P (2008) Evolution and the levels of selection Phil. Review 117(4):626–630
Forber P, Reisman K (2007) Can there be stochastic evolutionary causes? Philos Sci 74(5):616–627
Frank SA, Slatkin M (1990) Evolution in a variable environment. Am Nat 136(2):244
Gillespie JH (1974) Natural selection for within-generation variance in offspring number. Genetics 76(3):601–606
Glymour B (2011) Modeling environments: interactive causation and adaptations to environmental conditions. Philos Sci 78:448–471
Kant I (1998) Critique of pure reason. Cambridge University Press, Cambridge
Krantz D, Luce R, Suppes P, Tversky A (1971) Foundations of measurement (additive and polynomial representations), vol 1. Academic Press, San Diego
Lewens T (2004) The natures of selection. Br J Phil Sci 61:313–333
Lloyd EA (1988) The structure and confirmation of evolutionary theory. Princeton University Press, Princeton
Lynch M (2007) The origins of genome architecture. Sinauer Associates Inc., Sunderland
Matthen M (2009) Drift and "statistically abstractive explanation". Philos Sci 76:464–487
Matthen M (2010) What is drift? a response to Millstein, Skipper, and Dietrich. Philos Theory Biol 2:e102
Matthen M, Ariew A (2002) Two ways of thinking about fitness and natural selection. J Phil 99(2):55–83
Matthen M, Ariew A (2005) How to understand causal relations in natural selection: reply to Rosenberg and Bouchard. Biol Philos 20:355–364
Matthen M, Ariew A (2009) Selection and causation. Philos Sci 76:201–224
Mills SK, Beatty J (1979) The propensity interpretation of fitness. Philos Sci 46:263–286
Millstein R (2002) Are random drift and natural selection conceptually distinct? Biol Philos 17(1):33–53
Millstein R (2005) Selection vs. drift: a response to Brandon's reply. Biol Philos 20(1):171–175
Millstein R (2006) Natural selection as a population-level causal process. Br J Philos Sci 57(4):627–653
Millstein R (2008) Distinguishing drift and selection empirically: "the great snail debate" of the 1950s. J Hist Biol 41(2):339–367
Millstein R, Skipper RAJ, Dietrich MR (2009) (Mis)interpreting mathematical models: drift as a physical process. Philos Theory Biol 1:e002
Morrissey MB, Kruuk LE, Wilson A (2010) The danger of applying the breeder's equation in observational studies of natural populations. J Evol Biol 23(11):2277–2288
Morrissey MB, Parker DJ, Korsten P, Pemberton JM, Kruuk LE, Wilson AJ (2012) The prediction of adaptive evolution: empirical application of the secondary theorem of selection and comparison to the breeder's equation. Evolution 66(8):2399–2410
Northcott R (2010) Walsh on causes and evolution. Philos Sci 77:457–467
Okasha S (2006) Evolution and the levels of selection. Oxford University Press, Oxford
Otsuka J (2015) Using causal models to integrate proximate and ultimate causation. Biol Philos 30(1):19–37

Otsuka J (2016) Causal foundations of evolutionary genetics. Br J Philos Sci 67:247–269
Otsuka J, Turner T, Allen C, Lloyd EA (2011) Why the causal view of fitness survives. Philos Sci 78(2):209–224
Pearl J (2000) Causality: models, reasoning, and inference. Cambridge University Press, New York
Pence CH, Ramsey G (2013) A new foundation for the propensity interpretation of fitness. Br J Philos Sci 64:851–881
Pigliucci M, Kaplan J (2006) Making sense of evolution: the conceptual foundations of evolutionary biology. University of Chicago Press, Chicago
Ramsey G (2013) Organisms, traits, and population subdivisions: two arguments against the causal conception of fitness? Br J Philos Sci 64(3):589–608
Rausher MD (1992) The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. Evolution 46(3):616–626
Reisman K, Forber P (2005) Manipulation and the causes of evolution. Philos Sci 72:1113–1123
Rice SH (2004) Evolutionary theory: mathematical and conceptual foundations. Sinauer Associates Inc., Sunderland
Robb D, Heil J (2003) Mental causation. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/mental-causation/. Accessed 20 March 2016
Rosenberg A (1985) The structure of biological science. Cambridge University Press, Cambridge
Rosenberg A, Bouchard F (2005) Matthen and Ariew's obituary for fitness: reports of its death have been greatly exaggerated. Biol Philos 20(2–3):343–353
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 (1984) The nature of selection: evolutionary theory in philosophical focus. University of Chicago Press, Chicago
Sober E (1993) Philosophy of biology. Westview Press, Boulder
Sober E (2001) The two faces of fitness. In: Singh RS, Krimbas CB, Paul DB, Beatty J (eds) Thinking about evolution. Cambridge University Press, Cambridge, pp 309–321
Sober E (2013) Trait fitness is not a propensity, but fitness variation is. Stud Hist Phil Biol Biomed Sci 44(3):336–341
Spirtes P, Glymour C, Scheines R (2000) Causation, prediction, and search, 2nd edn. The MIT Press, Cambridge
Stephens C (2004) Selection, drift, and the "forces" of evolution. Philos Sci 71(4):550–570
Wade MJ, Goodnight CJ (1998) Perspective: the theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. Evolution 52(6):1537–1553
Wade MJ, Kalisz S (1990) The causes of natural selection. Evolution 44(8):1947–1955
Wagner GP (2010) The measurement theory of fitness. Evolution 64(5):1358–1376
Walsh D (2015) Variance, invariance and statistical explanation. Erkenntnis 80(3):469–489
Walsh DM (2007) The pomp of superfluous causes: the interpretation of evolutionary theory. Philos Sci 74:281–303
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):452–473
Woodward JB (2003) Making things happen: a theory of causal explanation. Oxford University Press, New York