

Natural Selection: A Case for the Counterfactual Approach

Philippe Huneman

Received: 6 December 2009 / Accepted: 30 June 2011 / Published online: 29 July 2011
© Springer Science+Business Media B.V. 2011

Abstract This paper investigates the conception of causation required in order to make sense of natural selection as a causal explanation of changes in traits or allele frequencies. It claims that under a counterfactual account of causation, natural selection is constituted by the causal relevance of traits and alleles to the variation in traits and alleles frequencies. The “statisticalist” view of selection (Walsh, Matthen, Ariew, Lewens) has shown that natural selection is not a cause superadded to the causal interactions between individual organisms. It also claimed that the only causation at work is those aggregated individual interactions, natural selection being only predictive and explanatory, but it is implicitly committed to a process-view of causation. I formulate a counterfactual construal of the causal statements underlying selectionist explanations, and show that they hold because of the reference they make to ecological reliable factors. Considering case studies, I argue that this counterfactual view of causal relevance proper to natural selection captures more salient features of evolutionary explanations than the statisticalist view, and especially makes sense of the difference between selection and drift. I eventually establish equivalence between causal relevance of traits and natural selection itself as a cause.

1 Introduction

It is usually said that natural selection “causes” adaptations, or “causes” the diversity in the biological realm, or above all causes evolution.¹ This paper aims at

¹ Among many instances: Ridley (1999, 90): “natural selection causes both evolutionary change and the evolution of adaptation”; Stern and Orgogozo (2008, 2155): “Natural selection causes predictable changes in phenotypic variation”; Nevo et al. (1997): “Natural selection causes microscale allozyme

P. Huneman (✉)

IHPST, CNRS/Université Paris I Sorbonne, 13 rue du Four, 75006 Paris, France
e-mail: huneman@wanadoo.fr

uncovering which conception of causation is needed to account for natural selection as a cause in those statements, and argues that a counterfactual view fits this role. Philosophical issues, such as “units of selection”, debates, or the debates over adaptationism (i.e. whether natural selection is the main cause of most of the most important traits) indeed involve a reference to selection as a cause. Such reference is also included in scientific controversies, such as the recent debates over “niche construction”.²

There are many kinds of explanations involving natural selection—be it in ecology, population genetics, behavioural ecology or molecular evolution studies. Especially, population genetics tracks the dynamics of selection and drift in a population of alleles, whose fitness values have already been assumed; whereas ecological investigations track what Wade and Kalisz (1990) call the “causes of selection”, namely the selective pressures which explain the fitness values, and behavioural ecology aims at capturing the relation between environmental demands and traits, or in other words, what the traits are adaptations for. Those are very different projects, yet both of them involve a concept of selection. The present paper is interested in a concept of natural selection at work both in population genetics and in ecologically oriented studies.

I take it for granted that the minimal explanandum of natural selection is the variation in allelic or traits frequencies (stability or change)—from which evolution, adaptation and diversity, and many of the things that natural selection is said to cause, can in part be derived; and this paper questions in which sense—if any—selection causes variation in the frequency of alleles or traits. In population genetics, natural selection since Fisher is modelled through a *statistical* construal; however many biologists, following Haldane, call “selection”, as well as drift and migration “*forces*”, that cause the evolutionary change. Natural selection, as a process occurring within a population of genes carried by individual organisms in ecological interactions, is therefore affected by some ambiguity. In this regard two viewpoints seem *prima facie* possible: first, we can say—apparently close to the original talk of many population geneticists—that selection is a genuine force or cause *acting on this population*; second, that we just have a complicated *set* of heterogeneous *individual interactions*, all occurring at individual level, and which results after some generations in the changing frequency of some traits and then some alleles, this population-level result being natural selection.

In the former perspective, selection, like drift and migration, is a force acting on populations that pushes them away from an initial hypothetical equilibrium

Footnote 1 continued

diversity”; Fox and Wolf (2006, 49): “Natural selection causes change in the phenotypic composition of a population through the differential birth and death of its members”.

² Odling-Smee et al. (2003) investigated the processes according to which activities of the organisms change the milieu, hence the selective pressures that bear upon them and supposedly shaped their traits. They claim that niche construction is “a neglected process in evolution” that has to be considered along with natural selection when explaining some organismal or ecological features. Yet one could object that niche construction is not another process than natural selection, it is only natural selection acting in a sophisticated manner; therefore, the whole debate relies on what one means by “causal process” when talking about selection and niche construction.

described by Hardy–Weinberg equations. This dynamical view³ requires that, in order to explain why an expected equilibrium is not reached in nature, we have to pinpoint some cause, and natural selection is the most important kind of causes used to explain such cases.⁴

The second, “statisticalist”, perspective, elaborated recently by Walsh, Lewens, Matthen and Ariew⁵ states that natural selection is not a cause but an upper level statistical outcome of individual causal processes. “Fitness and natural selection have no reality except as accumulations of more fundamental events” (Matthen and Ariew 2002). Saying that a trait has been selected is only a convenient way to talk of myriad individual interactions between organisms that involved those traits, and whose collective outcome is that on the average the relative frequency of the trait increased. The distribution of the traits and the interactions is enough to have the selective result, without requiring positing another cause superadded to them; natural selection instead appears as the statistical outcome of those interactions.⁶ So the main ontological point is to distinguish between genuine forces and their averaged population result: “Selection and drift are not forces that impinge on populations; they are statistical properties of a set of “trials”: death, births, and reproductions. The only genuine forces in evolution are forces that take place at the levels of individuals (or lower) and (...) none of them can be equated with selection or drift” (Walsh et al. 2002). Walsh (2007) refined this line of reasoning by arguing that selection or drift are only statistical properties in the sense that they depend on the set up used to describe the population, and their values vary in a correlative way according to how we choose this setting. The objective difference between selection and drift therefore disappears: both are the *same sampling process* described differently according to our knowledge of the fitness values and our partitioning of the population.

³ “Dynamical”, used first by Walsh, Matthen or Ariew, refers to the force talk used by population geneticists. Some authors prefer to talk of “causalists”, but this could be larger than the scheme described here. Nothing depends much upon these labels, anyway.

⁴ Stephens (2004) defends the force conception by analyzing this precise pattern of explanation.

⁵ Matthen and Ariew 2002; Walsh 2003.

⁶ Several arguments have been used by the statisticalists: the non-commensurability of selective pressures in a same fitness measure, which distinguishes selection from an addition of forces like in Newtonian mechanics ; the impossibility of distinguishing, among offspring in one generation and given the fitness values, those offspring that have been selected and those which are here by drift ; and therefore the non-commensurability of selection and drift which precludes one from treating evolutionary change as an addition of two forces. Objections have been proposed by Millstein (2005, 2006) considering selection as causal but taking place at a population level. Later Shapiro and Sober (2007) will use the same line of criticism as her: those who endorse a statistical view neglected the difference between “selection as an *outcome*” and “selection as a *process*”. As Shapiro and Sober (2007) say, “even if selection the product is not a cause of evolution of trait, selection the process might be”, but the statistical view seems to only consider “selection as an outcome”. However, the main point of the tenants of a statistical interpretation is that selection is *not* a causal process, so arguing against them about selection “as a process” seems to beg the question. Bouchard and Rosenberg (2004) answered by a defense of ecological fitness, Stephens (2004) rehabilitated the “force talk”, Brandon and Ramsey (2006) based their reply on considering how proceed selectionist explanations in ecological context, and Riesman and Forber (2005) used Woodward’s manipulationist view of causation to support a causal view of selection. Matthen and Ariew (2009), Walsh (2007) and Matthen (2009) refined the metaphysical and epistemological foundations of the statisticalist view.

There are more options than the dynamical and the statisticalist. Someone could accept the rejection of the first scheme, but resist the statisticalist view of selection as an aggregate *result*. It can be done either by arguing that selection is indeed not a force acting on individuals, but still, at the population level, a specific causal process involving population-level properties (such as frequency of traits) (Millstein 2006), or that, in the contrary, all causal interactions are at the level of ecologically interacting individual organisms but that this ecological setting is precisely the occurrence of selection as a causal process (Bouchard and Rosenberg 2004). However, this paper raises the question of whether any concept of causation would support the rejection of natural selection as a cause; therefore I am mostly concerned by the statisticalist position, which argues for such rejection. The position here developed will have, as a consequence, some affinities with the two other positions, but in order to only articulate it, I will be in what follows exclusively concerned by the statisticalist view.⁷

Given that causation allows for several metaphysical interpretations, it is indeed plausible that according to some conception there is a causal nature of selection, whether according to another one the statisticalists are perfectly right, so that the debaters often talk past, lacking of a common conception of causation. I first show that explanations by natural selection seem to involve relations of counterfactual dependence, hence causation in some sense, and present the relations between explanation and causation that can possibly hold about a simple analogous case. The second section specifies the relata of such counterfactual relation, answers some objections and describes how they behave in a specific biological case. The last section investigates the meaning according to which selection can therefore be said to be a cause, whereas most of the arguments by the statisticalists against a dynamical view of selection still hold.

2 Explanation and Causation: The Counterfactual Dependences Within Natural Selection

2.1 Natural Selection at Work: Forces, Causes and Interactions

Before starting, I notice that the initial papers about the statisticalist scheme mixed two distinct questions: Is selection a force? Is it a cause? It is easy to mention lots of

⁷ Glennan (2009) has developed a view very similar to the present one. Especially, like in this paper, he relies on a discussion between two meanings of causation, “relevance” and “production”—while I distinguish “difference-making” or counterfactual aspect of causation, from the “process” aspect, and then we apply this difference to the analysis of selection’s causal nature (see Hall (2004) for a defense of this duality in meaning). Similarly, he accepts that, as to production, there is nothing more than the individual interactions, so that selection is not a cause in this sense; but that in terms of causal relevance there is something causal about natural selection, which is not exhausted by the analysis of causal (*sensu* production) interactions at the level of individuals. The difference between our views is in the way I understand this causal relevance, and the counterfactual analysis I make about the connection between trait values and trait frequencies. However, the bulk of the present argument has been written before its publication, so I did not know the paper and owe the knowledge of its existence to an anonymous referee. Given that it’s a case of independent convergence of views, I acknowledge here the proximity and will not detail at length the points of agreement, which would be useless.

causes that are not forces (“I could not meet you because I missed the train”), so the two issues are obviously separate. In the dynamical conception selection is a force, for sure, but the major point actually is the explanatory pattern used by Sober (1984) to describe evolutionary biology. The force-talk is somehow analogical, first of all because the context-dependency of natural selection⁸ prevents us from making it a force on a par with genuine physical forces like electromagnetism or gravity, which are defined by properties independent of their context (represented by variables: mass, distance, charge). Yet we could agree with Sober that selection is not a force but works *like* a force in explanations.⁹ The remaining question is then: how to assess the value of the causal talk used by evolutionary biologists about natural selection modelled in statistical terms by population geneticists? The statisticalist answer grants it no value.

Millstein et al. (2009) objected to the statisticalists that they draw conclusions about natural selection on the basis of considering one population genetics model, but models as mathematical construal are neutral regarding causal interpretations. However, even if the statisticalists considered only a Fisher-Wright model of population genetics, whereas other models exist and hold different assumptions about drift, still the set of population genetic models has a minimal overlap, upon which the arguments of the statisticalists bear, and which constrains all causal interpretations: drift is therein statistically construed as a stochastic effect depending upon population size, which does not allow one to specify the subset of the set of after-evolution-resulting alleles due to drift and the subset due to selection. Against their objection, Matthen (2009) considers a Moran model of population genetics as an example of what he called “statistical abstraction”, which is the way population genetics models in general proceed.

In making their case the advocates of the statistical view clearly distinguished two things: selection is not a *cause* but it *explains* (because the genuine causes are on the individual level of the trials of life, whose statistical average is collected by us in the predictive “fitness values”). They took for granted that explanation and causation are different terms and that they could be separated (Walsh (forthcoming)). But how to understand this distinction between *to cause* and *to explain*, and precisely in this case: how is causation to be thought in order to allow for cases where explaining does not amount to causing? The statisticalists’ argument presupposes that all causal *relations* are causal *interactions and processes*. It is only

⁸ This word means that the selective pressures which are at work and determine which traits are likely to be selected will always depend upon the ecological framework.

⁹ The issue of force had been considered by Endler (1986); his analogon was erosion, and he forcefully rebutted the force-talk. More recently Lewens (2009) has shown the exact measure within which Sober says selection is a force; he distinguishes then the “force of selection”, which can be equated with variation in fitness and does not exactly allow one to capture what causes evolutionary change, and “selection-for property”, which is causal but may actually depart from the force of selection and its predictions. As Lewens acknowledges, this detailed understanding of Sober’s view considerably decreases the difference with the statisticalist view, because all of them agree that what actually causes evolutionary change, at the ecological level (“selective forces” *sensu* Lewens (2004)), differs from the “force of selection”, captured in population genetics models. However, the contrast between a dynamical and a statistical view of selection is still useful to deal with the problem, and corresponds to the contrasts between statistical and causal properties advocated by Walsh (2007).

in this sense that no superadded cause or “tertium quid” (Matthen and Ariew 2009) is needed to understand evolutionary changes, once you have all the individual causal interactions. They are committed to a view of causation akin to Salmon’s process-view of causation (1998),¹⁰ which allows them to say that since *all* causal interactions take place at the level of individual organisms, natural selection as a population-level phenomenon cannot be causation.¹¹ On the basis of this commitment they can claim that all the individual interactions *cause* an evolutionary event, but that the statistical properties involved in natural selection just *explain* it.

However, besides Salmon’s process-view of causation, there are actually conceptions of causation based on a “*difference-making*” scheme (Menzies 2001; see Waters (2007) for an application to genic causation): A causes B means that A makes a difference on whether (or how) B occurs. Difference-making is clearly instantiated by the counterfactual view of causation elaborated by Lewis (1973—see now Hall and Paul 2004), which basically says that “A causes B” means “B would not be here if A had not been here”¹²—but the probabilistic conceptions of causation also rely on such scheme (the purported cause makes a difference on the probability of the effect). I argue in this paper that the causal character of natural selection appears when one holds a conception of causation based on the scheme of difference-making. A counterfactual instantiation of this scheme will be the most appropriate then.

Let’s consider a very ordinary case of established selection, in order to establish that conceiving of selection as a cause in counterfactual terms is highly plausible. A species of *Poeciliid* fish has conspicuous colours; this is understood as the result of selection (Endler 1982). This explanation could be explicated in this way: because some fish are of this colour rather than another colour, the fish from this subclass reproduce more than others, and the overrepresentation of such colour leading to the fixation of the colour in the species has as a cause the colour itself, because it is closer to the one of the background algae, and hence prevents their carriers from being seen by predators. *Prima facie*, the state of the population—in terms of frequency of colour properties—is thereby counterfactually dependent upon the fact that the property of having this colour is available to individuals in the population. Besides, Endler notices that in another population from an allopatric *Poeciliid* fish species, the predators are not the same but the same colour is selected. So, interestingly, although selection does not cover here the *same class of individual interactions* (since it is not the members of the same class of predators that are the

¹⁰ Walsh (2000) uses the famous process/pseudo-process advanced by Salmon (1998) and defined in terms of transmitted marks in order to enunciate his thesis, that selection is not a cause but a “pseudo-process”.

¹¹ This commitment to local, process-based causation has been recently made clear by Matthen and Ariew (2009)—discussed in a companion paper to this one.

¹² The counterfactualist view is more complex, because Lewis elaborates upon this concept of counterfactual dependence a concept of causation which excludes from all counterfactual relations these which would lead to backtracking counterfactuals, problems of preemption, etc. These subtleties are not relevant for my point here, so I skip them; I will later on take from Lewis’ analysis what is necessary to make my case regarding selection.

selective agent); however, it is the same *fact* (having the colour property) which is causally relevant.¹³

Another example illustrates those properties of convergent evolution in regard to my question. Snails are mostly dextral rather than sinistral, regardless of habitat. Foraging specialisation in dextral rather than sinistral snails may then be advantageous, and it has been demonstrated that asymmetry in the snail-feeding apparatus has evolved in some aquatic arthropods. However, Hosoi et al. (2007) have shown that it's also the case for some terrestrial vertebrates, namely the snake *Pareas iwasakii*. Those snakes extract the soft body of snails from the shell because their jaws are not powerful enough to crunch it. Since most of the snails are dextral, it is likely that the “snake predators may improve the efficiency of snail body extraction by differential action of the left and the right mandibles” (169). Therefore, authors checked three facts: non snail eating specialists of parent species have a symmetrical dentition; sinistral morphs of snails are longer to be handled by snakes than dextral forms; and consequently “*pareas iwasakii* failed in the predation of sinistral snails more frequently than in the predation of dextral” (170). This makes an argument for the asymmetry of dentition being caused by natural selection. When the authors write “the dentition asymmetry should therefore be an adaptation for improved performance in the extraction of the dextral soft body” (ibid.) they summarize those facts.¹⁴ Since in the aquatic arthropods asymmetry has been selected for snail eating, we have a “convergent directional asymmetry” (169). Clearly being asymmetric in dentition makes a difference about the number of snails that can be preyed on. Such difference implies difference in differential reproduction. If their dentition were not asymmetrical those snakes would clearly not reproduce as they do, so the trait “asymmetric dentition” here is actually the cause of its own increase in frequency, then its fixation, through the increase in frequency of snakes with asymmetric dentition. Therefore, biologists can say—in a statement whose legitimacy will be established in the following—that natural selection caused the arising of asymmetric dentition in *both* those snakes and the aquatic arthropods eating dextral snails, because the *same* kind of property (mandible asymmetry) is causally relevant in an analogous context (preying habits). Yet the statisticalist view cannot make sense of this causal commonality of natural selection across genera and contexts, because in this perspective, the individual interactions underpinning these cases of selection are of two different kinds. The causal processes underlying the fixation of the asymmetry are indeed different in the two cases; what is common is only the kind of result, namely the arising of sinistral dentition, but not the cause, whereas it is natural for a biologist to see here a common cause at stake.

¹³ Endler (1986) gave an early critique of the idea of selection as a force, arguing for a view of selection very close to the one articulated later by the statisticalists. In this case, the statisticalist analysis of course fits the case of the *Poeciliid* no less than it fits the other cases of selection; however, I emphasize a consequence of this study, namely the possibility of building higher level causal patterns of selection, to which the general statisticalist view seems *prima facie* inappropriate.

¹⁴ The authors note that this asymmetry is important enough to be a “key innovation” since it “initiated the adaptive radiation of pareatine snakes through south east Asia as dextral snails eaters” (170).

Now, suppose indeed that one wants to exclude the hypothesis of drift when explaining those colours of *Poeciliid* fish. If colour red were there because of drift, the statement telling it could be formulated in the following terms: “if their colour had been neutral, they might still have been likely to increase in frequency”. This opposes indeed the counterfactual “if the colour had not been conspicuous, then this colour would not have been increasing in frequency”. Such contrast makes attractive the idea that counterfactuals express causation by natural selection—because they express the fact that in selection, *as opposed to drift*, the fact that the increasing trait *is what it is* (and therefore has the effects it has) is causally relevant in its increasing frequency. Hence a counterfactual framework to make sense of natural selection as a cause naturally accounts for the distinction biologists make between selection and drift—whereas we saw that for the statisticalists this difference is not essentially and conceptually objective.

However, isn't this view that there is population-level causation (frequency is clearly a population-level fact) only an appearance, given the critiques raised by the statisticalists, namely that all causes are to be found at the level of the interactions between all fish, and that no causes should be superadded to that? Talking of natural selection is then just an explanation, and allows predictions based on fitness values, and my causal reading of natural selection would here only pertain to explanation. But in the present picture, we are committed to another relation between explanation and causation: causation can occur at both levels, provided that one does not use it in the same sense: it would be processes at the lower level, and counterfactual dependence at the higher level. In order to justify this general scheme, so that counterfactual causation would not fall under the objection that it posits unwanted epiphenomenal causal relations, I now explore a simple analogous case. Then in the next Sections 3 and 4 I will specify the counterfactualist account of natural selection as a cause, and deal with some of its difficulties.

2.2 Causation and Explanation at Two Levels: A Simple Analogon to Biological Evolution, Involving Counterfactual Causation

The general issue raised by the case of natural selection is the relation between a set of individual causes and what is explanatory at the level of a population.

“Explanation” is actually a pragmatic concept or, at least, the pragmatic dimension is irreducible if you deal with what explaining is.¹⁵ An explanation, basically, relates two statements. It has contrast classes: “p explains q” means that p makes us understand why q,¹⁶ which has always as an implicature “why is it q rather than q’?”.¹⁷ The choice of “rather than q’” instead of “rather than q” pertains to pragmatics. Pragmatics here is only understood in terms of explanatory

¹⁵ Achinstein (1979) All these preliminary considerations on explanations are controversial but since this paper does not address explanations in general, I have to rely on some extant theories to sketch a difference between explanation and causation.

¹⁶ In fact, some explanations are not answers to why questions, as when we make clear what we said earlier, etc. But about natural selection expected explanations are naturally likely to be casted into why questions (why does pandas have opposable thumbs?, etc.).

¹⁷ Dretske 1972; Sober 1984.

needs and strategies, for sure. And what makes x explanatory in a given explanation, i.e. the fact that x causes y , is not depending upon our explanatory interest; what is pragmatic here is only the fact that in this explanation I will consider x and in that other explanation, I will consider x' , x and x' being each as well equally objectively connected to the explanandum. Causation is less pragmatic: it relates facts or events. But when A causes B , this unique causal relation entails several explanatory relations, since each contrast class (B_i , contrasting with B) calls for a distinct explanation.

In many cases, the *explanatory* and the *causal* relations are indeed not at the same level, as the statisticalists claim it is the case with natural selection. Think of a *realization relation*, like in the famous example of the cylinder peg and the hole by Putnam (1975): the cylinder's molecules there realizing the shape, which occurs at the level of macroscopic entities: the shape explains what the molecules cause (the peg not entering the hole). Think also of what physicists call *attractors*: they *explain* the set of trajectories but don't *cause* it, like the slope of a mountain: all the stones rolling along this slope will come down to the valley. This valley can be called an attractor of the stones; unless we wish to explain something about one particular stone, the best explanation will be the attractor. However, the attractor has no causal power by itself, since what causes the stone to come down in the valley is its own causal pathway and, as a general cause, the force of gravity acting at each step of its course. Pinpointing attractors explains, in a way that abstracts away from a variety of causal stories. Granted, you might need some causal information to explain why X is an attractor, but X explaining some trajectories of a system is not by itself and entirely a causal story.¹⁸

Clearly however, natural selection does not involve a *realization* relation: both types of traits competing are not as a whole realising some property that would concern trait distribution. It does not either imply an attractor: the fate of one carrier of a given trait, unlike the fate of a stone on a hill, is not explained by a final state included in the final distribution of traits. So, are causation and explanation really disjoined in the case of natural selection, as claimed by the statisticalists?

A simpler example of a process where interactions between individuals are the *causes*, and resulting population level trends are *explanatory*, would be the case of the crumpled \$1 bills: if you look at a \$1 bill in your wallet, it is very likely that it will be more crumpled than a \$100 bill. Why are \$1 bills on the average more crumpled than \$100 bills? Each \$1 bill undergoes a monetary circuit made of exchanges between economic agents. In the end of the circuit the bill is crumpled in a specific way. This sequence of exchanges is the *cause* of its crumpled state. But the *nature* of the singular exchanges in the process is not a cause since the same final state of crumpledness might occur with different singular exchanges: suppose

¹⁸ Arguing for this claim would start with the fact that a demonstration of X being an attractor relies on some mathematical properties of X as an object in the phase space. The argument for the detachment of causal stories and attractor explanations is that two different systems might have the same or similar attractors X and X' as mathematical objects in the phase space, which will be both attractors because they display the same mathematical properties, so they are attractors for the same reasons; yet those reasons are not causal stories, because, the systems being of different natures, the causal stories they involve are wholly distinct.

that exchange no. 2 and exchange no. 3 are switched, or that exchange no. 2 is replaced by a very different transaction, the result will be the same. What is causally relevant here is the *frequency* of exchanges. This implies that, since a \$100 bill is less likely to be used when giving back change, the \$100 bill is less likely to be exchanged the same number of times, so in the end a set of \$1 bills will be more crumpled than a set of \$100 bills.¹⁹ This fact is not dependent upon the nature of the singular exchanges which constitute each monetary circuit undergone by a bill.²⁰

Clearly, the monetary circuit it underwent *caused* the state of one \$1 bill, and the sum of these circuits caused the general state of crumpledness of the \$1 bills, as compared to a less crumpled state of \$100 bills. However, if one of those \$1 bills had undergone other monetary exchanges, or if n bills had been two times more exchanged and n other bills exchanged two times less, the result would have been the same. So if one wants to *explain* the general state of the \$1 bills (as compared to the \$100 bills), one has not to enter into the details of the monetary circuit of those bills, but can say that on the average \$1 bills have been more exchanged than \$100 bills. This perfectly valid explanation, in turn, is *explained* by the fact that \$1 bills, being less valued, are more likely to be given when one gives change, so are more likely to undergo much more exchanges, therefore to be more crumpled after a given period of time than \$100 bills.

So, like in the case of natural selection according to the statistical view, the addition of individual interactions is *causally* relevant (the “trials of trade”, could we say in reference to Walsh et al.’s “trials of life”); and what is caused is at the level of a population (of bills) a general trend (\$1 bills tend to be more crumpled). The ultimate *explanation* is here the value of the bills. But “being worth \$1” does not belong to the chain of monetary transactions that causes the state of crumpling of one bill; hence *it is not part of the causal process*.²¹

Yet we know that *frequency* is the causally relevant parameter (rather than the nature of each singular exchange), and that exchanges of \$1 bill are so frequent *because* it is a \$1 bill rather than a \$100 bill, which means that if the bill were not a \$1 bill there would not be such a pattern of exchanges (more precisely, the pattern of exchanges would not fall in the class of patterns having such a high frequency of exchanges). Therefore, it would not be more crumpled than a \$100 bill. So we have a counterfactual dependence between “worth \$1” and “more folded than a \$100

¹⁹ Of course, one must take here into account the background condition constituted by the fact that currency is divided into bills of \$1, \$5, \$10 etc. If you change this distribution, if bills are of \$1, \$2, \$3 and \$100, the results about crumpledness will be different, however the same reasoning about the explanatory role of the value of the bills will still hold.

²⁰ I do not enter here into complications proper to the real cases in a given economy, since the US Federal Reserve System can decide to create immensely more \$1 bills than \$100 bills, which would oppose my demonstration. I will not either consider the fact, which is true, that bills that are too crumpled are eventually replaced (but they have to be really very crumpled). My example is a bit like examples in combinatorial theory, when one first treats the simple cases of urns with no replacement of balls, before considering the with-replacement case. Nevertheless, the empirical appearance of our various bills in real life seems very close to the ideal case I consider, making plausible that the two facts I just mentioned are in real life as negligible as they are neglected in my ideal case.

²¹ First of all because no conserved quantity is transmitted between the economical property “\$1” and one given exchange—while the process-view of causation requires conserved quantities.

bill”, thereby being an instance of the property “worth \$1” causes the bill to be “likely to be more crumpled than a \$100 bill”. Hence, provided a counterfactual account of causation, the property of being worth \$1 *causes* the fact that on the average the bill is more folded than a \$100 bill (because of the difference between \$1 and \$100). So, to say it in a sentence: “*Because the \$1 bill is worth \$1*, it enters a chain of heterogeneous individual interactions whose result is that it will probably be more crumpled than other bills.” The loose “because of” employed here can be rigorously cashed out in terms of counterfactual causation, if we apply Lewis’s definition: amongst worlds where the token \$1 bills are worth much more (\$100), then the closest ones are such that those bills are now not much more crumpled than the \$100 bills—otherwise, it seems that we should have to make many changes to the economic system in general to get the same pattern of crumpledness between all individual bills.²² Therefore, if in such context of a two-levels phenomenon, some properties can be causal in a counterfactual sense and cause population-level properties, it is plausible that natural selection could be both an *aggregate result* of individual-level events causing a population-level trend (i.e. the variation in allelic frequency), an *explanation* of such trend—as claimed by the statisticalists—and could also embody *causal relationships* involving properties which cause the trend. The following section will elaborate the proper counterfactual account of selection as a cause and explains why it holds.

3 Specifying the Counterfactual View and Answering Objections

3.1 Drift and the Problems with Counterfactual Reasoning: What are the Relevant Counterfactual Dependencies?

The previous description sounded like all *Poeciliid* fish of some color had more reproductive success than others because of this color; yet it is not always the case in one population that all organisms *as* carrying the trait *a* which ultimately will be evolutionarily successful have been fertile because of *a*. It might be that for example some *bs* got unlucky, struck by lightning as in the usual examples, etc.: suppose this gazelle is indeed a fast runner, but it had no need to compete with its slow rival, which got struck by lightning... This might not change anything about biologists saying “natural selection has occurred”, but it would make our counterfactual statement invalid. The problem here is that in the unlucky gazelle case, the counterfactual supposed to express natural selection is not true.²³

²² In fact, we should have to allow very different values for things, and another set of bills which would range up to something like \$10,000, which seems further from our world than simply multiplying the \$1 bills in order to keep the structure \$1/5/10/50/100.

²³ Another example would be the last individual in a small population of mimetic insects; all the good ones (good camouflages) have been eaten, there is only this (low fitness) one, and it gets eaten. The counterfactual “if it had been of the other color, it would have survived”, intuitively expressing the fact that it has been selected against, won’t hold (well, it’s the last insect left, it probably would have been eaten anyway).

So suppose that you have a small population. Drift occurs, meaning that you have higher chances that some “random” events, i.e. rare events that have no reasons related to the focal trait to happen, will affect the organisms (add gamete sampling, which is not rare but is essentially indiscriminate). It might be that the faster gazelle will be at the boundary of a territory and gets no resources: its being slow or fast doesn’t impinge on its survival. If the population were large, lots of fast gazelles in the middle of the territory will compensate this lost gazelle, so the unlucky gazelles wouldn’t mean much for intergeneration change of frequencies. Hence, the problem is that in a small population our counterfactual characterization of selection and drift is just untrue: the population being small, the nature of the traits has fewer chances to be causally relevant across all the interactions. This is not to say that there is no counterfactual dependency definitive of selection in small populations, but rather that we have to identify which counterfactual dependencies are in general relevant. Counterfactuals bearing *on individual organisms* are not valid here, precisely because of small populations. In the extant populations, it is not the case that for any slow gazelle x , “if it had been faster it would have survived”, and for a collection of gazelles sharing the same speed g , it is not the case that “if trait “running at speed g ” was “running at $i < g$ ”, it would have not gone to fixation”: suppose that we have only very few individuals running at $n > g$ mph and they all fall into a hole...²⁴ So the domain on which the counterfactual is valid is not the individuals in extant populations but *a set of hypothetical populations* with varying size and the same repartition of traits. In this case, it is true that “if fast gazelles had not been so fast they would not have outreproduced slow gazelles in most of the hypothetical populations”.²⁵

More formally, suppose now a simple model where you have fast and slow gazelles (speed being a variable with two possible values: $X = (a, a')$), wolves, and random death affecting indifferently both types of gazelles. Then assume that we have an infinity of those populations, with uniformly varying sizes, and each having quite the same initial ratio of a/a' , and we let them evolve. It is not always true for a gazelle a of speed a which survives and leaves two offspring (in population H) that it did it *because* of its being faster (suppose that it competed once with another fast gazelle that randomly died, etc.). But the fact that type A *more frequently* wins over type A' (in a series of populations H_i) is clearly due to the fact that A gazelles run faster. So if all those gazelles *were not* the A type, they would have won in less cases than they did. Hence the fact that type A is *more likely* to win (i.e. *ceteris paribus* leave more offspring in more H_i) than type A' counterfactually depends on type A being the faster type.²⁶ With this formulation, the cases of fast gazelles A faring worse than slow gazelles A' for various reasons, most probably in small

²⁴ Lewens (2009), developing Sober’s account, see cases of selection for the low fitness variants as example of a divorce between selection as a force (defined by variation in fitness), and selection for properties (being actually the cause of evolutionary success). However, here there is no “selection for properties”, for reasons explained in the next paragraph.

²⁵ Computer simulations can often indicate the behavior of a set of hypothetical populations, because we can keep fixed the parameters (size, initial frequencies) and run the dynamics many times.

²⁶ If one accepts objective chances (Mellor 1995), then this formulation could easily be turned into chances talk, chances being the hypothetical limit frequencies.

populations, and the cases of fast individuals A faring better than slow gazelles A' *for reasons independent of speed*, will not falsify the counterfactual, since those few cases won't alter the frequencies.²⁷

In the following, the counterfactual statements about increase in frequency are implicitly all understood as statements enunciated in a set of hypothetical populations, as defined in the previous paragraph. Then, what does this counterfactual relation I am considering actually relate? Take a running gazelle: to say that there is selection for speed (V) is to say that increase in frequency of gazelles running at a high speed v counterfactually depends on the value of V. So more generally, if we represent a trait by a variable X, taking the values x_1, x_2 , etc., which are the possible variants (it can be the many alleles at a locus, in the framework of population genetics; the current formulation is supposed to explicate the concept of natural selection in a way adequate both for models in population genetics and for explanations in ecology or paleontology), the counterfactual dependency underwritten by natural selection holds between (a) "the fact that for some entities (organisms or genes or whatever...) X takes value x_i ", and (b) "the increase in relative frequency of instances of organisms (or entities) whose value of X is x_i "²⁸. The case of quantitative genetics is self-evident because the value of variable X is just the value of the trait considered: speed, size, etc. Population genetics just requires discrete values of the variable; the reasoning is in both cases the same.

We can then enunciate "selection for x_i " as a causal statement in the following terms: "If their value of X had not been x_i , those entities would not have been likely to outreproduce other entities with other values of X (hence) x_i would not have increased in frequency"²⁹. The last consequence in the statement ("hence etc.") is ensured by the heritability of X, which means that among organisms having the same value x_i of X, a significant proportion of their offspring will still be x_i , or in case of continuous values, the distance of the average of their values of X to the average value of X will remain close to the value of such distance for the parent organisms. In the following, I assume that X is always heritable enough; reproductive success of a subclass of organisms and increase in frequency of a value of X shared by those organisms therefore are equivalent. And of course, this characterization of selection directly translates into models of population genetics, where selection acts upon alleles or genotypes: values of X are rival alleles (or alternative genotypes) determining *ceteris paribus* the value of the trait.

Talking in terms of variables allows one to handle cases where traits are hypothetical rather than extant variants: it suffices to consider values not actually taken. It subsequently enables one to consider cases of negative or purifying selection where some variants are stable against continually arising new variants (by mutation), optimality models or ESS models where we compute the reproductive values of possible variants and strategies, or models where fitness is defined in terms

²⁷ This consideration solves the "last living insect" case too.

²⁸ This increase, more rigorously, has to be understood in the context of sets of hypothetical populations as noted above.

²⁹ Notice also that those counterfactuals are not backtracking because an evolution by drift is always possible.

of resistance to invasion (adaptive dynamics). Besides, throughout this paper I have equated selection with increase in relative frequency—but the counterfactual statement is easily modifiable in order to account for cases of stabilizing selection where *frequency* of the trait will not change but *variance* of the distribution of traits diminishes. It may suffice to formulate the causal statement in terms of counterfactual dependency between the value of the x_i and the decrease in relative frequency of the x_j with $j \neq i$. Stabilizing selection for x_i goes like this: “if the value of X for those entities had not been x_i , then instances of the other many x_j would not have been likely to decrease in frequency.”

Manipulationist accounts of causation are built on a counterfactual account of causation, provided that the notion of “intervention” defines which counterfactual changes are made upon the value of the variables (Woodward 2003). Riesman and Forber (2005) argued for an interpretation of selection as a cause based on such view; clearly their analysis would be compatible with my view, and generally with any interpretation of the causal status of selection that would rely on the scheme of difference-making. However, I still prefer the counterfactual interpretation, because the manipulationist view of selection as a cause is vulnerable to specific objections.³⁰ Now I still have to show how and why in the case of natural selection these counterfactual dependencies actually hold, and this will explain why they make selection differ from drift.

3.2 Possible Worlds: Reliable Factors, Causes and Selection

A counterfactual assertion is, as Lewis pointed out, an assertion about some possible worlds and their distance from our actual world. Selectionist explanations, when phrased in counterfactual terms, should indeed embody references to those possible worlds, and I will show this now. Recall first that not all counterfactuals in biological contexts express natural selection. If an individual organism is struck by lightning, the counterfactual “if it had not been here it would have not been killed” holds, yet it is not about natural selection. Why? Because lightning strikes are very *rare* circumstances, and hence their effects appear only in *small* populations. Once again, lightning or any case of random deletion can kill all the high fitness individuals if the population is small, but probably not if it is large. Suppose you have A and A' types, and their fitness is $W_a > W_{a'}$.³¹ So it can be the case that the frequency of A' becomes higher than the frequency of A because of those random deletions, but it will probably not be the case in a large population. And if this indeed obtains (for example if lightning often suppresses lots of A s) then we will

³⁰ Lewens (2009) and Walsh (forthcoming) actually emphasized cases of Gillespie fitness, i.e. cases where defining fitness includes the variance of the trait distribution, and hence changing fitness implies changing population size, which makes impossible to distinguish an intervention on fitness (hence selection) and an intervention on population size (hence drift), so that no legitimate intervention, hence no causation in the manipulationist sense, can be defined.

³¹ This is a rather crude model, not even a Fisher-Wright model for single locus diploid population, since we consider only genotypes and not alleles; all the subtleties of population genetics are here not at stake, it's like a population of asexual haploids. However, the conceptual point made here still concerns real population genetics, because as we saw the values of the variables in the counterfactual statement can be understood as competing alleles in a one locus model.

realize that we initially made a mistake, so we have to count lightning as a selective pressure and then reconsider the fitnesses. In effect, if the *As* fare worse than the *A's* regarding the lightning, which seem in fact frequent, those *As* are actually less fit than *A's*. This shows that *we define a selective pressure as a factor reliably occurring in a population and preferentially affecting one type rather than another*. But in a very small population, it is difficult to identify those selective pressures, because it is difficult to know what is reliably occurring and makes a difference regarding the types through its occurrence; in this case, it will be hard to recognize what is due to selection, and what is not, concerning the change in traits/alleles frequencies.

We call drift the effect of those non-reliably occurring factors—which are all the more effective when the population is small. On this basis, drift and selection are conceptually distinct (Millstein 2002). To use the distinction made by Millstein, here the addition of those random factors is “drift the process”; their added effect is “drift the outcome”.

Defining selective pressures as those reliably occurring factors which differentially affect the trait types makes clear why some counterfactual dependences occur, which characterize selection and not drift. Suppose that in a large population in its environment we identified the reliably occurring factors which affect preferentially one trait, and the chance factors.³² Now, the *closest* possible worlds where the fact that traits are what they are in the actual world, and the same organisms have them, are the worlds *W* in which we change the non-reliable circumstances, such as lightning striking here rather than there (even if lightning is biologically unsatisfying and imaginary made up as an instance of drift...). A possible world *W'* where a *reliable* circumstance is not there—for example, wolves don't eat gazelles—is further from ours than the latter. A selected trait *a* is a trait such that it fares better than its variants with regard to the selective pressures, which are defined by those reliable circumstances. It is then likely that if those organisms had not this trait *a* that they have—which defines them as “organisms having trait *a*”—, in the closest worlds *W* with such change the differential reproductive success of those organisms would not occur, because precisely what may change in those worlds is only the non-reliably occurring factors, the ones that are not involved in selection, hence nothing will compensate the fact that those organisms fare differently (or: are likely to fare differently, or on the average fare differently...) regarding those factors because of their not having trait *a*. If one wants to compensate such fact, one should consider a very far possible world *W'*, because the reliably occurring factors would have to be changed. For the same reason, if a trait *a* of organisms *A* gets fixed by drift, it is not true that in the closest possible worlds where organisms *A* don't have this trait they don't enjoy their differential reproductive success—since the chance events leading to the fixation of *a* can happen even if the organisms *A* are *b*-organisms not displaying *a*, because there is no correlation between those events and *a* (unlike selective pressures and *a* in the case of selection), so that we don't

³² Note that if there are several traits, it is highly probable that any reliably occurring factors will at least differentially affect one type of trait so in practice, reliably occurring factors suffice to define selective pressures.

have to do lots of other changes in the actual world to get a close possible world W where those counterfactual b -organisms face the same chance events than organisms A faced, and fare equally well. In order to make those b -organisms *counterfactually* have the same reproductive output that they *actually* have as organisms A , the only thing we have to change concerns some non-reliably occurring factors, so it will happen in worlds quite close to ours.

The present section talked in terms of organisms carrying traits because it was easier to grasp—however, the formulation that we previously had, in terms of values x_i of the variable X representing the trait, is here straightforwardly deducible. Indeed, if drift is responsible for the fixation of a trait x_i , it is the case that in the closest world where you change the value of the instances of this trait from x_i to x_j —for instance change all the green eyes into blue eyes—the x_j still goes to fixation, because in this world the chance events that constituted drift still occur.³³ Inversely, if x_i has been selected, it is the case that in the closest worlds where organisms are x_j instead of x_i then x_j won't get fixed because it will fare worst than its competitors regarding the reliably occurring factors that do still occur in those worlds since they are the closest ones.

Let's apply this to our example of colored fish. Suppose that the *Poeciliid* fish are of their actual color because of drift. This means that, in the closest possible worlds where *Poeciliid* fish have another color, the events leading to their reproductive success would still occur, which is true: this color would still have gone to fixation, because the events responsible for its fixation are those non-reliably occurring events that do still occur in the closest worlds without *Poeciliid* of the actual color.

3.3 Selection in the Field: A Case of Frequency-Dependence

This refined analysis of selection as a case of counterfactual causation can be applied to a more sophisticated case of natural selection in the field: asymmetrical predation in cichlid fish *Peridossus microlepis* from Lake Tanganyika in south-east Africa. This is a clear case of frequency-dependent selection. Of course, many cases of natural selection are in principle frequency-dependent; things here are like non linearity in physics, where lots of cases display non linear terms, though most of the time they can be neglected.

Let's explore the case: fish of the seven species from the genus *Peridossus* in Lake Tanganyika display a lateral asymmetry of the mouth opening, because “an individual's mouth opens either rightward or leftward as a result of an asymmetrical joint of the jaw to the sensorium.” (Hori 1993, 217) A study has been led on the population of *Peridossus microlepis*. Those fish eat scales on the flank of bigger fish; researchers showed with a study on live lures that right handed (dextral) individuals always attack their victim's left flank and reciprocally; this was confirmed by checking the content of dextral and sinistral *Peridossus*, which was only full of scales of the other's orientation prey. This is putatively explained by the

³³ Here one can define drift with Millstein (2006) in terms of biological processes like gamete sampling, etc.; the point is that in various possible worlds as close as one wants, those events won't produce the same outcome, unlike what we call selection.

fact that the distorted mouth enlarges the area of teeth in contact with the prey's flank, but only when it's facing the corresponding flank: "thus, the correspondence between the handedness and the side of attack should be a functional requisite for the success in feeding of these scale eaters" (ibid.). It is also a fact that handedness is heritable. In this context, it is plausible to think of handedness of mouth as a result of natural selection.

But obviously it is predictable that "when one of the two phenotypes—for example dextral individuals—is more abundant in the population, prey fishes will tend to guard more against attacks to their left side, which results in sinistral individuals gaining greater success and a commensurate greater fitness" (218) and then the frequency will get back to equilibrium. The field study has indeed shown support for this balancing mechanism: the frequency of both types along 11 years oscillates around an equilibrium (Fig. 1) and the hypothesized mechanism of frequency dependent selection is likely to be the cause since it is established that "when sinistral individuals were numerically dominant the prey suffered scale eating from dextral individuals more frequently than sinistral individuals and vice versa" (Hori 1993, 218). Moreover, the molecular analysis of the group makes it plausible that the *Peridossus* species evolved their asymmetrical morphology by acquiring a scale eating habit from older carnivorous species: "recurrent evolution of similar feeding morphology appears to have occurred with specialization to scale eating" (Takahashi et al. 2007, 196). To this extent, asymmetry in jaws seems likely to be an adaptation proper to this group.

Let's interpret such explanation of the balancing frequencies of types of *Peridossus microlepis*. Suppose we are at a moment (point M in the diagram) where there is a majority of dextral fish. This subpopulation, along with the rest of the environment (which is generally stable there),³⁴ constitutes the environment of the subpopulation of sinistral fish. Concerning them, we can say that if they were not left jawed, they would less easily reach their preys (which are more accustomed to watch their right side) and then would not reproduce as they do. I write it this way: variable H (handedness) has two values D and S. At this moment the increase in frequency of S will counterfactually depend upon S itself. Let's now formulate the causal efficiency of selection here as a general formula, in order to justify its phrasing in terms of counterfactuals. The specificity of this frequency-dependent selection mechanism (one of the few known cases which rests on prey-predator relations) consists in integrating the frequency of S and D individuals in measuring the counterfactual dependency of such frequencies upon the nature of S and D. Writing s (resp. d) = Freq. S (resp. D), if " $d \gg s$ ", then "if S individuals had not been S then they would have not enjoyed differential reproductive success and S would not increase in frequency". The frequency d defines the closest worlds according to which we evaluate the counterfactual dependence of frequency increase in S individuals upon the trait S. In effect, worlds where d is close to its actual value are closer than the actual one (everything else being equal) than worlds where d is far smaller. Suppose then that we check the closest worlds where a fish

³⁴ "Habitat conditions in this area are very stable and most resident fish populations are highly persistent with respect to their density" (Mori 1993, 217).

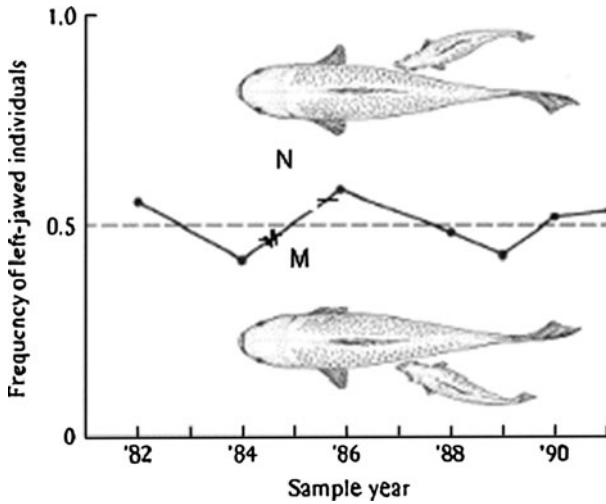


Fig. 1 Frequency-dependent selection on *Peridossus microlepis* (Hori 1993)

actually instantiating S is not S . In those worlds, this fish would be D , hence it would probably have far smaller preying success (because of the disadvantage for commonness, and because D fish are common, since d is close to its actual value, so it's high), therefore less reproductive success than it is likely to have in the actual world. Possible worlds where such a fish is likely to have high reproductive success are further than those ones since they need to have a smaller value for d . The same reasoning holds when $s \gg d$ and when we consider the selection for dextral fish.³⁵ In the appropriate (counterfactual) sense of causation, there is therefore a causal interpretation of the way frequency dependent selection explains the alternative increasing of frequencies.

3.4 A Caveat: Ontology and Epistemology of Selection

To sum up this section, the strategy was threefold. First, I have shown that there are counterfactual dependences involved when one speaks of natural selection, so these are good candidates to be the meaning of causation when one speaks of natural selection as a cause. Then I investigated what these counterfactual dependencies could be, and between what relata they could hold; finally I have shown why they do hold, in terms of possible worlds (but any non-actualist modal metaphysics such as Stalnaker's also do the job, indeed),³⁶ and applied this to a field example. However, this last step concerned only simplified models of natural selection. What I have shown is that, if there is natural selection, and if it is taken to be causal, this causation is captured by the determinate counterfactual dependencies, and such

³⁵ This value d will also define the variable to manipulate in the Reisman and Forber's account.

³⁶ Some could find Lewis's modal actualism a high price to get some causal meaning of selection, but nothing in the present argument is tied to this interpretation of possible worlds; Stalnaker also made sense of counterfactuals with possible worlds but with a modally weaker metaphysics.

dependencies indeed hold when you cash out the counterfactuals, because of the involvement of “reliable factors” in natural selection. It does *not* imply that these conditions are sufficient for natural selection, i.e. that, each time you have traits (in a trait space) carried by individuals, then you have a selective dynamics. No one could show this, because it’s a fact of population genetics that, given the conditions for selection (heritability, variability and fitness), you do not necessarily have evolution by natural selection (lots of epistasis, or dominance, or pleiotropy, or population size, can preclude selection).

The remaining worry here would be about the fact that in reality many traits are simultaneously under selection, so besides X you have many correlated variables here, making the evaluation the counterfactuals difficult overall. This has been the source of many worries by philosophers, which can be traced back to issues of the content of functional ascriptions (Neander 1990; Dretske 1986, etc.). For example, suppose that, changing the value of X, race speed, also increases the value of Y, which is race-speed-related attractiveness by mates, or Z, the sexual investment potential (negatively correlated to race speed for metabolic reasons). Suppose we are considering a case where biology says that X evolved by natural selection. It might be the case, then, that if the value of X were different, still X-carrying organisms (who are also carrying a specific value of Y or Z) would *not* decrease in frequency. It seems that natural selection and counterfactuals therefore part company, or at least, that it’s difficult to test the counterfactual I associated with natural selection.

One has to be clear here about the scope of the question. With only X and a population of organisms, I showed which causal relations could hold that can count as natural selection. This was at the level of conceptual analysis; yet the issue of how reality falls under the concept (or not) is different. In other words, the epistemic issue of assessing counterfactuals in realistic cases where many traits are simultaneously under selection is an issue for the methodology of evolutionary theory, and evolutionary theory itself. Up to some point, the counterfactuals can be supported by interspecies comparisons, for example, of sister species in different environments (Harvey and Pagel 1991) or non-sister species in parent environments (Endler 1986). And then, if you have three correlated traits X, Y and Z, and you consider the population dynamics, the mathematical theory gives you tools to estimate the selection on the correlated traits (Lande and Arnold 1983). So in this case, given that you get quantitative relations in the form of regression coefficients, you can sort out the contributions of selection on each correlated trait, which can be in return indicative of the counterfactual relations between traits and their change in frequencies. The point here is that once we show ontologically that counterfactual dependencies hold in virtue of reliability of factors in environment, and that these dependencies clearly, in very simple cases, underpin selection ascriptions about the evolution of trait X, there is room for the scientific question of how, in more complicated real cases, those counterfactual dependencies definitive of natural selection can be tracked. I just emphasized here the distinction between ontological issue about the concept of selection as a cause, and epistemological issues about disentangling selective hypotheses, which have been already richly investigated.

The next section assumes that these epistemological issues are solved, so that the objections based on the variety of traits under selection can be left aside.

4 Causal Relations at Several Levels: Natural Selection as a Cause and the Causal Relevance of Traits

We just established that natural selection explanations are causal explanations—in the sense of specifying a counterfactual dependency between a trait and its increase in frequency (roughly speaking). However, here we have not yet proven that natural selection is *a cause* and the position defended here is not distinguished from the statisticalist claim—indeed, it is quite compatible with it.

The preceding discussion has established the statement

(S): the fact that organisms Y_x have trait X is causally relevant in their success in differential reproduction, and then in the increase of the frequency of X

I agree with the statisticalists that selection is not a cause as a “tertium quid”, another causal process superadded to all the causal interactions between individuals: the reasons for which selection is not so are there compelling. But when they say that the only causes are the individual interactions occurring between individual organisms, they equate the causal working of selection with the *actual* interactions between organisms. It overlooks the fact that some other interactions between individuals may lead to the same frequency of traits—exactly like in the \$1 case, where the *particular* monetary exchanges *as such* don’t necessary make a difference regarding the outcome, because what is relevant is the frequency of those interactions. Concerning selection, suppose that I have a population of individuals with two types of traits (A and B), and such that their interactions lead to some frequency of trait a relative to trait b . Now if I switch those A individuals which, due to some interactions with B individuals, had no offspring, with other A individuals, it will *not make any difference* to frequencies of offspring types, granted that the total number of individuals are the same (before selection). So all the sets of individual interactions which have the same frequencies of interactions between A and B with a same given outcome, are equivalent regarding the difference they make concerning type frequencies. To this extent, it is not the collection of individual interactions which *causes* (*sensu* causes as difference makers, at least) the change of gene frequencies, but only the *pattern of distribution* of trait a and trait b which is instantiated by this particular collection and *yet could have been instantiated otherwise*. What is causally relevant is not the individual interactions—“shark x eats fish y ” etc.—but their addition *in so far* as it instantiates a peculiar distribution involving the property under selection, and this distribution can be instantiated by a huge variety of individual interactions. So even if there is no causal process over and above the individual interactions, what cause the variation in gene frequencies are not the individual interactions *as such*. The counterfactual criterion used above would not obtain if applied to them. What causes (*sensu* counterfactual dependence) the variation—evolution, stability, etc.—is indeed the fact that their collection belongs to a *class of patterns of interactions* that are equivalent regarding

any substitution of traits and individuals as considered above: “either shark a_i eats fish F, or shark a_j eats fish F, etc.”, each a_i instantiating property A. Such class, written $C(A)$, is correlated to property A. What causes evolution is a *disjunction* of possible collections of interactions between individual organisms (in $C(A)$). All distributions of traits allowing for a pattern of interactions in $C(A)$ therefore cause evolution.

But, exactly like in the case of the bill where the disjunction of possible monetary circuits for a \$1 bill and its financial value were correlated, this disjunction of collections of interactions is such that all its members satisfy the proposition “the a sharks have property A rather than another shark property”, and were they not to have this property the pattern of interactions would not yield the same differential output: this thereby makes property A causally relevant here. It is because sharks a_i display property A, and because indeed A is A (with all its effects) that any member of $C(A)$, the class of distribution of interactions between individual organisms, results in *change in traits frequencies* in the next generations. So if this property A is causally relevant in evolution, then the distribution of interactions in $C(A)$ which is actually instantiated results in the frequency change.

If we write (S') “*the set of interactions causes the change in frequencies of traits*”—a statement to which the statisticalists would agree, as write Matthen and Ariew (2009, 202)—, then (S) (the causal relevance of property A in evolution) entails (S') (evolution being caused by the fact that one of the patterns of interactions in $C(A)$ is instantiated). And reciprocally, if a distribution of traits giving rise to some set of interactions causes evolution, it is the case that some trait in this distribution is causally relevant in evolution. Otherwise, you could change all traits with no causal consequences; but you would not have the same distribution of traits, hence the same interactions and then evolutionary change, which is self-contradictory. Hence (S') reciprocally entails (S), thus (S) and (S') are equivalent, one being stated in terms of aggregated individual interactions, the other in terms of counterfactual dependence, and both describe the same phenomenon.

Now, if causation is not only causal process but also counterfactual dependence, when we say (Q): “*selection causes change in the frequency of traits*”, we are not committed to saying that selection is a causal interaction or process over and above *individual* interactions. To this extent, the *most natural meaning* of (Q)—precisely because the “*tertium quid*” meaning is rejected—is that a causal relation occurs in the form stated above, namely a causal relevance of some trait, that is (S) a counterfactual dependency of the increase³⁷ in frequency of such trait upon the fact of some organisms having this trait. If biologists talk of selection as a cause, and if, when selection occurs, there is one causal relationship occurring, which is indeed the counterfactual dependence I have highlighted, then the meaning of this causal talk can only be such causal relevance. (Q) “*selection causes evolution of trait a* ” cannot mean anything but (S). So (Q) is indeed logically equivalent to the fact that (S') the causal interactions between individuals result in the average increase of

³⁷ The account is refined above (Sect. 2) in order to take into account stabilizing selection, which could be objected to the present formulation.

some trait, *since* this fact is *equivalent* to the causal relevance of selected traits in their own frequency increase (S).

There is here one last ambiguity left: what exactly is causal, and are we speaking of “causation” or “causal relevance”? In fact (S) talks of *causal relevance* of traits, whereas (Q) talks of *selection* as “a *cause*”, and they proved to be equivalent. It is crucial here to distinguish between degrees of generality: the causal relevance of a trait in its increase in frequency is an instance of a very general causal scheme, which is the dependence of the population-level fate of traits (alleles) upon their nature and effects. To say that there is “selection for x” means exactly such causal relevance of trait x. “Natural selection” is, on the other hand, the causal scheme under which all cases of “selection for x” fall: to say that “natural selection” occurs means that there is (at least) some x such that there has been selection for x; hence that such causal relevance holds.³⁸ So in such case, it is true that “natural selection” is the causal scheme responsible of the changing frequency of the traits, so it is indeed a cause of such fact. But the emphasis on the causal talk will depend on the context, i.e. on which effect is considered (e.g. asking whether a trait is an adaptation for some environmental demand, vs. testing a genomic sequence for the signature of selection).

5 Conclusion

I claim that selection is both causally and explanatory relevant (rather than being shorthand for infinitely numerous individual causal processes). This commits me to a counterfactual view of causation, relating facts and properties rather than events. Since causation is not to be thought as causal process, the distinction between processes and pseudo-processes (Salmon 1998; Walsh 2001), essential in the process account of causation, is of no help to understand natural selection. Although selection is not a force—except by a remote analogy—the statistical view is not wholly accurate for several reasons, mostly because the elementary properties of natural selection cannot be thought as state variables like entropy, which are solely predictors. I insisted on the counterfactual view of causation because otherwise, if one holds a process view of causation, one cannot resist the statistical conclusion about selection, namely that the only causal properties are at the level of individual interactions between organisms, because there are actually no other processes—as Matthen and Ariew (2009) rightly asserted. The statisticalist view is indeed right in many aspects: the inseparability of selection and drift at the individual level, the impossibility of considering selection as a force or superadded causation over and

³⁸ Lewens (2009) pointed out that in some cases “selection for” cannot be equated with the force of selection, based on the variations in fitness values. Yet “natural selection” here is not what he calls “force of selection”, but just the causal scheme subsuming the specific explanations by natural selection for properties. The full account of this difference needs another paper aiming at distinguishing how selection differs from drift in each viewpoint. At the most general level, here I agree with Lewens’ conclusion that defining the nature of selection cannot avoid to specify the explanatory context conditioning which adaptations occurred, and especially to decide whether one explains rather the ecological context or the variations in allele frequencies.

above individual interactions... Its only weakness is that it assumes a conception of causation that compels their authors to find no legitimate meaning to statement (Q) “natural selection causes evolution”, against its commonsensical understanding. Such understanding is legitimized as soon as one acknowledges non-interaction and non-process causal relations, because it acquires a natural meaning in terms of (S): “having trait *a* has been causally relevant in the increase of instances of trait *a*”, which means also “there has been selection for *a*”. This view allows one to save a sharp distinction between selection and drift. Thereby, those debates are not of pure metaphysics, but involve consequences for the science.

Acknowledgments The materials from this paper have been presented at the CSHPS meeting in York (2006) and the University of Toronto workshop on the statistical interpretation of selection (2008). The author is grateful to the audiences for their suggestions and critiques. Many thanks in particular to Jean Gayon, Jim Griesemer, Françoise Longy, Francesca Merlin, Roberta Millstein, Michael Strevens, Denis Walsh, for their useful comments on earlier versions. The text has been carefully language-checked by Denis Walsh, whom I warmly thank for that. Finally, I thank two anonymous referees for having greatly improved the paper.

References

- Achinstein, P. (1979). *The nature of explanation*. New York: Oxford University Press.
- Bouchard, F., & Rosenberg, A. (2004). Fitness, probability and the principles of natural selection. *British Journal for Philosophy of Science*, 55, 693–712.
- Brandon, R., & Ramsey, G. (2006). What’s wrong with the emergentist statistical interpretation of natural selection and random drift. In M. Ruse & D. Hull (Eds.), *The Cambridge companion to philosophy of biology*. Cambridge: Cambridge University Press.
- Dretske, F. (1972). Contrastive statements. *Philosophical Review*, 81(4), 411–437.
- Dretske, F. (1986). Misrepresentation. In R. Bogdan (Ed.), *Belief, form, content and function*. Oxford: Clarendon Press.
- Endler, J. (1982). Convergent and divergent effects of natural selection on color patterns in two fish faunas. *Evolution*, 36, 178–188.
- Endler, J. (1986). *Natural selection in the wild*. Princeton: Princeton University Press.
- Fox, C., & Wolf, J. (2006). *Evolutionary genetics: Concepts and case studies*. Oxford: Oxford University Press.
- Glennan, S. (2009). Productivity, relevance and natural selection. *Biology & Philosophy*, 24(3), 325–340.
- Hall, N. (2004). Two concepts of cause. In C. J. Hall & L. A. Paul (Eds.), *Causation and counterfactuals* (pp. 225–276). Cambridge: MIT Press.
- Hall Ned, C. J., & Paul, L. A. (Eds.). (2004). *Causation and counterfactuals*. Cambridge: MIT Press.
- Harvey, P. H., & Pagel, M. D. (1991). *The comparative method in evolutionary biology*. Oxford: Oxford University Press.
- Hori, M. (1993). Frequency-dependent natural selection in the handedness of scale-eating cichlid fish. *Science*, 260, 216–219.
- Hoso, M., Asami, T., & Hori, M. (2007). Right-handed snakes: convergent evolution of asymmetry for functional specialization. *Biological Letters*, 3(2), 169–173.
- Lande, R., & Arnold, S. (1983). The measurement of selection on correlated characters. *Evolution*, 37(6), 1210–1226.
- Lewens, T. (2004). *Organisms and artefacts. Design in nature and elsewhere*. Cambridge: Cambridge University Press.
- Lewens, T. (2009). The natures of selection. *British Journal of Philosophical Science*, 1–21.
- Lewis, D. (1973). *Causation*. New York: Oxford University Press. Philosophical papers II.
- Matthen, M. (2009). Drift and ‘Statistically abstractive explanations’. *Philosophy of Science*, 76, 464–487.

- Matthen, M., & Ariew, A. (2002). Two ways of thinking about natural selection. *Journal of Philosophy*, 49(2), 55–83.
- Matthen, M., & Ariew, A. (2009). Selection and causation. *Philosophy of Science*, 76, 201–224.
- Mellor, D. H. (1995). *The facts of causation*. London: Routledge.
- Menzies, P. (2001). Difference-making in context. In C. J. Hall Ned & L. A. Paul (Eds.), *Causation and counterfactuals*. Cambridge: MIT Press.
- Millstein, R. (2002). Are random drift and natural selection conceptually distinct? *Biology and Philosophy*, 17(1), 33–53.
- Millstein, R. (2006). Natural selection as a population-level causal process. *British Journal for the Philosophy of Science*, 57, 627–653.
- Millstein R., Skipper R., & Dietrich M. (2009). (Mis)interpreting Mathematical Models: Drift As A Physical Process. *Philosophy and Theory in Biology*, 1.
- Neander, K. (1990). Dretske's Innate Modesty. *Australasian Journal of Philosophy*, 74(2), 258–274.
- Nevo, E., Apelbaum-Elkahert, I., Garty, J., & Avigdor, B. (1997). Natural selection causes microscale allozyme diversity in wild barley and a lichen at 'Evolution Canyon', Mt. Carmel, Israel. *Heredity*, 78, 373–382.
- Odling-Smee, J., Laland, K., & Feldman, M. (2003). *Niche-construction: The neglected process in evolution*. Princeton: University of Princeton Press.
- Putnam, H. (1975). The meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science*, 7, 215–271.
- Ridley, M. (1999). *Evolution*. Oxford: Oxford University Press.
- Riesman, K., & Forber, P. (2005). Manipulation and the causes of evolution. *Philosophy of Science*, 72, 1113–1123.
- Salmon, W. (1998). *An At-At theory of causal influence. Causality and explanation*. New-York: Oxford University Press.
- Shapiro, L., & Sober, E. (2007). Epiphenomenalism—The Do's and the Don'ts. In G. Wolters & P. Machamer (Eds.), *Studies in causality: Historical and contemporary*. Pittsburgh: University of Pittsburgh Press.
- Sober, E. (1984). *The nature of selection*. Cambridge: MIT Press.
- Stephens, C. (2004). Selection, drift and the "forces" of evolution. *Philosophy of Science*, 71, 550–570.
- Stern, D., & Orgogozo, V. (2008). The loci of evolution: how predictable is genetic evolution? *Evolution*, 62(9), 2155–2177.
- Takahashi, R., Watanabe, K., & Nishida, Hori. M. (2007). Evolution of feeding specialization in Tanganyikan scale-eating cichlids: a molecular phylogenetic approach. *BMC Evolutionary Biology*, 7, 195.
- Wade S., & Kaisz S. (1990). The causes of natural selection. *Evolution*, 44(8), 1947–1955.
- Walsh, D. (2000). Chasing shadows—Natural selection and adaptation. *Studies in the History and Philosophy of Biology and the Biomedical Sciences*, 31, 135–153.
- Walsh, D. (2003). Fit and diversity: Explaining adaptive evolution. *Philosophy of Science*, 70, 280–301.
- Walsh, D. (2007). The pomp of superfluous causes: The interpretation of evolutionary theory. *Philosophy of Science*, 74, 281–303.
- Walsh, D. (forthcoming). Variance explanations.
- Walsh, D., Lewens, T., & Ariew, A. (2002). Trials of life: Natural selection and random drift. *Philosophy of Science*, 69, 452–473.
- Waters, C. K. (2007). Causes that make a difference. *Journal of Philosophy*, 104, 551–579.
- Woodward, J. (2003). *Making things happen*. New York: Oxford University Press.