Studies in History and Philosophy of Biological and Biomedical Sciences 44 (2013) 336–341

Contents lists available at SciVerse ScienceDirect

Studies in History and Philosophy of Biological and Biomedical Sciences

journal homepage: www.elsevier.com/locate/shpsc

Trait fitness is not a propensity, but fitness variation is

Elliott Sober

Philosophy Department, University of Wisconsin, Madison, WI 53706, USA

ARTICLE INFO

Article history:
Received 17 June 2012
Received in revised form 25 October 2012
Available online 13 April 2013

Keywords:
Extensionality
Fitness
Optimality
Propensity

ABSTRACT

The propensity interpretation of fitness draws on the propensity interpretation of probability, but advocates of the former have not attended sufficiently to problems with the latter. The causal power of C to bring about E is not well-represented by the conditional probability $Pr(E|C)$. Since the viability fitness of trait T is the conditional probability Pr(organism O survives to adulthood|O has T), the viability fitness of the trait does not represent the degree to which having the trait causally promotes surviving. The same point holds for fertility fitness. This failure of trait fitness to capture causal role can also be seen in the fact that coextensive traits must have the same fitness values even if one of them promotes survival and the other is neutral or deleterious. Although the fitness of a trait does not represent the trait's causal power to promote survival and reproduction, variation in fitness in a population causally promotes change in trait frequencies; in this sense, fitness variation is a population-level propensity.

The propensity interpretation of fitness began with Brandon (1978) and Mills and Beatty (1979), but it subsequently evolved, so that philosophers now sometimes differ in how they understand what the view says. Here is how I will understand it in what follows: First, there is the important insight that individuals of identical fitness can differ in how successful they are at surviving and reproducing. The individuals have the same *abilities*, but good luck for some and bad luck for others can lead to unequal outcomes. The abilities in question need to be described probabilistically. An individual's viability fitness is its probability of surviving from egg to adult. Its fertility is some function of the different probabilities the individual has of producing different numbers of offspring once it has reached adulthood.[1] These probabilistic abilities are called *propensities*, with a nod to the propensity interpretation of probability. Finally, the fitnesses of token individuals get linked to the fitnesses of traits by a simple formula: the fitness of a trait is the average fitness of the individuals that have that trait. Just as the fitness of an individual is a propensity, so the fitness of a trait is a propensity; it represents how having the trait affects an individual's survival and/or reproductive success.

The starting moves in this progression of ideas are on the right track, but things go wrong when it comes to trait fitnesses. The fitnesses of token individuals may be propensities, but evolutionary biology has little use for that kind of fitness. Rather, it is the fitnesses of traits that matter to the science. Can those fitnesses be interpreted as propensities? I argue in what follows that they cannot. Reasons for this negative conclusion can be found by thinking about the propensity interpretation of probability in the light of Humphrey's (1985) objection; complementary reasons may also be found in the fact that coextensive traits must have the same fitness value even if one of them promotes survival and reproduction while the other is neutral or deleterious (Sober, 1984, 1993).[2]

In what follows, I first discuss the connection between the fitnesses of token individuals and the fitnesses of traits. I then take up the question of what role the fitnesses of token individuals play in evolutionary biology; I argue that they are useless, since

E-mail address: ersober@wisc.edu

[1] Brandon (1978) and Mills & Beatty (1979) recommend equating fertility fitness with expected number of offspring (= the mathematical expectation), but problems for this proposal, including those described by Gillespie (1977), were recognized by Beatty & Finsen (1989). Brandon (1990) argues that the propensity interpretation need not be wedded to defining fertility fitness as expected number of offspring. Sober (2001), Pence & Ramsey (2013) discuss some of the problem cases.

[2] Drouet & Merlin (in press) argue that fitnesses are not propensities in the sense of the propensity interpretation of probability, but their arguments are different from the ones I propose. They also defend two claims about explanation: (i) fitness differences among token organisms do not explain why they differ in reproductive success; (ii) the explanatory power of trait fitnesses derives from the Law of Large Numbers.

1369-8486/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.shpsc.2013.03.002

their values typically cannot be estimated. It is trait fitnesses that are important to the science. I then argue that the fitness of a trait does not represent the trait's propensity to promote survival and reproduction. If trait fitnesses don't represent causal powers, where are causal descriptors of evolution by natural selection to be found? I describe two—the concept of selection-for (Sober, 1984) and the variation in fitness that a population exhibits (Millstein, 2006). Population-level variation in fitness is a causal propensity.

1. How individual and trait fitness are related

The distinction between the fitness of a trait and the fitness of a token organism is widely recognized, as is the following connection between them (Brandon, 1978; Mills & Beatty, 1979; Sober, 1984):

(1) The fitness of trait T in population P and environment E at time t = the (arithmetic) average fitness of the organisms that have T in P and E at t.[3]

Although some may see this proposition as an arbitrary stipulation, in fact it follows from two reasonable assumptions:

- that the fitness of an organism in its environment is identical with the fitness of the total trait complex that the organism possesses in its environment.
- that certain probabilities are identified with actual frequencies.

For simplicity, consider the component of fitness that pertains to viability. Organism Os viability fitness (w_v) is its probability of surviving the passage from egg to adult. The first bulleted claim means that Os viability fitness can be expressed as a conditional probability

(2) $w_v(O) = Pr(O$ survives to adulthood$|O$ at the egg stage has $T_o)$.

Here T_o is the totality of traits that O has at the egg stage. T_o is a *total trait complex*, whereas the T mentioned in (1) is *any* trait (e.g., the "singleton" trait of having a heart).

If there are n organisms ($O_1, O_2, \ldots, O_n$) that have trait T in the population, the axioms of probability entail that

(3) $w_v(T) = Pr(O$ survives to adulthood$|O$ at the egg stage has $T) = \sum_{i=1}^{i=n} Pr$ (O_i survives to adulthood$|O_i$ at the egg stage has $T_i)Pr(O_i$ has $T_i|O_i$ has $T)$.

Proposition (1) follows from (2) and (3) if $Pr(O_i$ has $T_i|O_i$ has $T)$ equals the actual frequency of each total trait complex among the n individuals that have T. If each of those n individuals has a unique total trait complex, then the actual frequency of each trait complex is $\frac{1}{n}$.

Proposition (2) is motivated by the general philosophical idea that an organism's fitness in an environment supervenes on the totality of traits that the organism possesses in that environment.[4] Once the total trait complex and the environment are specified, the organism's fitness is fixed; there is no "primitive thisness" (*haecceity*) that has any relevance. Proposition (3) is a truth of mathematics, but it does not say how the value of $Pr(O_i$ has $T_i|O_i$ has $T)$ should be understood. Why should its value be the actual frequency of T_i in T? Actual frequency is the wrong interpretation for many of the probabilities used in science. Why should it be used for this one? The reason is that one wants the fitness of a trait to help predict what will happen to the n flesh-and-blood individuals that have the trait.

A principle of extensionality follows from Proposition (1): *If traits A and B attach to exactly the same individuals in a population, then the two traits have the same fitness value in that population* (Sober, 1984, 1993). Extensionality, a concept from the philosophy of language, is a property of linguistic contexts. If the sentence "the fitness value of trait A in population P is x" is true, then "the fitness value of trait B in population P is x" will also be true if A and B are coextensive in P. Coextensive terms (terms that apply to exactly the same objects) may be substituted one for the other in extensional contexts, *salva veritate*.

Advocates of the propensity interpretation of fitness have argued that fitness must not be equated with actual survival or reproductive success, and I agree. This may suggest that the probabilities that are used to define fitness must never equate probability with actual frequency. This suggestion is mistaken. Interpreters of probability are often pluralists, thinking that probability sometimes has a subjective meaning (representing rational degree of belief) while at other times it means something objective. Even so, they sometimes shy away from using different objective interpretations of probability for the different concepts that occur within a single theory, perhaps because they are concerned to avoid the fallacy of equivocation. The fear is exaggerated; mixing is routine in models of evolution where some probabilities represent actual frequencies while others do not.

Proposition (1) may suggest that a trait's fitness value can be determined by first discovering the fitness values of the organisms possessing the trait and then taking the average. This is not what Proposition (1) says, which is fortunate, since the suggested epistemology is mistaken, as I now will explain.

2. Individual fitnesses are useless

It is often remarked that evolutionary biology takes no interest in the fitnesses of token organisms. For example, biologists don't bother with the fitness of Charlie the Tuna, though they may want to discuss the fitness of tuna dorsal fins (Sober, 1984). This is not an arbitrary decision. The fitnesses of token organisms are typically inaccessible.[5] This is because organisms taste of life but once. Consider a population of coins that are physically different in which each coin is tossed a single time. The outcome of the single toss of a given coin provides too meager a resource for estimating what the probability was that it would land heads. In contrast, the frequency of heads in this population may be a reasonable estimate of the average probability of landing heads that a coin has in this population. If organisms lived their lives again and again, their viability (and fertility) fitnesses would be more accessible. If cats had nine lives, that would be a start.

Can individual fitness values be inferred by looking at multiple organisms that are physically identical and that live in identical environments? Supervenience entails that they must have the same fitness. With enough of these carbon copies, we could estimate the viability fitness of each by looking at the frequencies with which these carbon copies survive to adulthood. Perhaps each of the unique elephants in a herd could be cloned 1000 times, and 1000 replicates of the herd could be formed, each placed into an identical environment. The fitnesses of each token elephant in the parent herd could then be estimated by watching what happens to its 1000 replicates. The sense in which this protocol is

[3] Henceforth, I will usually suppress mention of population, environment, and time, which should be taken as understood.
[4] It is a good question how "the environment" should be specified. I leave that question aside in my discussion of Proposition (1).
[5] I say "typically" because there are exceptions. For example, if you know that a trait is lethal in childhood, you can assign a newborn that has that trait a viability fitness of zero.

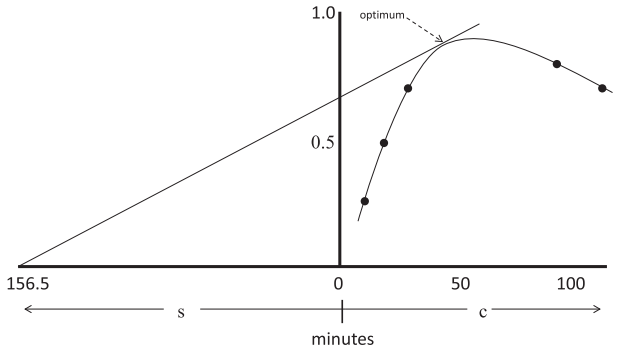

Fig. 1. Parker's (1978) optimality model of male dung fly copulation time.

"possible" is entirely disconnected from what is biologically feasible. Biologists are almost always stuck with the unique organisms they observe. Token fitness is a pipe dream.[6] Mills and Beatty's (1979, p. 276) statement that "it is the fitness of types which figures primarily in explanations of microevolutionary change" is an understatement.

Ramsey (2006, pp. 492–493) takes issue with the thesis that the fitnesses of token individuals are unknowable by proposing the following analogy. One way to tell what a coin's probability is of landing heads is by examining its physical make-up and the physical make-up of the tossing device used. Another is to toss the coin multiple times or to toss multiple coins "of the same type." He says "we can do the same thing with organisms." My reply is that maybe we can, *in principle*; but, *in practice*, this isn't workable at all. A coin's probability of landing heads on a given toss is a simple function of a few of its physical properties and a few of the properties of the tossing environment. An egg's probability of surviving to adulthood is vastly more complex. And I've already commented on the fact that the total trait complexes of token organisms are almost always unique. The upshot is not that fitnesses are unknowable, but that the fitnesses of token organisms are.[7] Biologists get along quite well by thinking about the fitnesses of traits that are shared among organisms; the fitness of the total trait complex that a token organism uniquely possesses is typically beyond their ken.

We see here a peculiar inversion of a standard epistemic complaint against Platonism—that knowledge of particulars is easy or unproblematic, whereas knowledge of properties is hard or impossible (see, e.g., Benacceraf, 1973). This purported asymmetry flips like a rabbit/duck when knowledge of probabilities is considered. When token organisms are unique and live just once, their probabilities of surviving from egg to adult are usually impossible to estimate, but the viabilities of the traits they share are often easy to estimate.[8]

Do well-motivated optimality models render the fitnesses of token organisms accessible? They do not. For example, in Parker's (1978) model of male dung fly copulation time, the optimal time for a male to spend in copula is derived from assumptions about (i) the time a male spends searching for a mate and then guarding her after copulation, and (ii) the proportion of eggs that will be fertilized, conditional on the time spent in copula. See Fig. 1. The *x*-axis represents time. If a male spends *s* minutes searching and guarding, what value of *c* is optimal, where *c* is the amount of time spent copulating? The *y*-axis represents the proportion of eggs that are fertilized; its value depends on copulation time. The curve relating *c* to *y* is estimated from data; notice that it represents diminishing returns. The value of *c* that is optimal is the one that maximizes the proportion of eggs fertilized *per unit time*. The slope of the diagonal line in the figure represents this maximal rate. Parker did not estimate the value of *s* from data on just one male; rather, he pooled males and worked with the average amount of time they spend searching and guarding; likewise, the curve representing *y* as a function of *c* was based on data from multiple parental pairs. What Parker derived was the optimal value of *c* for *the average male.* But even if he had worked with a single male, the point of importance would be the same: this model does not say what the fitness value is of any individual or trait. It merely entails a *fitness ordering* of trait values; there is an intermediate optimum with fitness declining monotonically with distance from that optimum. This ordering suffices to make some predictions; fitness point values aren't needed, which is fortunate. Optmality models are about traits, not Charlie the Tuna.

3. Is trait fitness a propensity?

To answer this question, we need to review what a propensity is and how the propensity interpretation of fitness relates to the propensity interpretation of probability (see Galavotti, 2005; Gillies, 2000; Hájek 2011 for discussion of the latter).

> Propensities are causal powers. When we say that water-solubility is a dispositional property that sugar has, we don't just mean that
>
> If sugar were immersed in water (in the right way[9]), it would dissolve.

We also have in mind two causal claims:

> Immersing sugar in water (in the right way) would cause the sugar to dissolve.
> Sugar has a property that causes it to dissolve when it is immersed (in the right way).

When sugar dissolves, this is caused both by the environment it is in and by its internal make-up.

This causal dimension in the concept of disposition recurs in the idea of probabilistic propensity.[10] According to the propensity interpretation of probability, "$Pr(A|B) = p$" means that p is the strength of the causal tendency that B has to bring about A.[11] This causal idea has long been part of philosophical discussion of the propensity interpretation; it explains why Humphreys' (1985) criticism of that interpretation is so widely regarded as decisive. Sometimes $Pr(A|B)$ is high even though B is the effect and A is the cause. If your probability of having smoked cigarettes is high, conditional on your now having lung cancer, this doesn't mean that having cancer later caus-

[6] My objection here is epistemic, not semantic. I do not claim that these single-case probabilities "make no sense," just that scientists rarely are able to make good estimates of their values.

[7] My point about token organisms is not just that we usually can't be *certain* what their fitness values are; my claim is that we rarely are able to have *good estimates* of those values.

[8] For more on Platonism and evolutionary theory, see Sober (1981, 2011b, §4.7).

[9] This phrase is present to exclude immersing a lump of sugar after encasing it in wax, for example.

[10] See, for example, Peirce (1910), Popper (1959, 1990), Giere (1973), Fetzer (1982), Miller (1994), though Gillies (2000, pp. 114, 133) takes the unusual step of defining "propensity" to mean any non-frequency-based objective probability; he is okay with "noncausal propensities."

[11] Here it must be noted that long-run propensity theorists do not want to say that a fair die has a weak propensity (of $\frac{1}{6}$) to land 3 on the next roll, but prefer to say that the die has a strong propensity to land 3 with a long-run relative frequency of approximately $\frac{1}{6}$. Short-run propensity theorists do not need to choose.

ally promotes smoking earlier. The propensity interpretation is mistaken as a thesis about *some* objective probabilities.[12,13]

Humphreys' counterexample suggests others. For example, $Pr(A|A) = 1$ so long as A is not a contradiction, but that doesn't mean that every noncontradictory proposition has a maximally strong causal tendency to make itself true. But the counterexample that is most relevant to the propensity interpretation of fitness is the following: *even when B causally promotes A, the conditional probability Pr(A|B) is usually a poor measure of Bs causal power to bring about A.* Suppose that studying philosophy in your 20s has a mild positive effect on your chance of reaching age 50. This is perfectly consistent with *Pr*(you live at least 50 years|you study philosophy in your 20s) being very high. Causal powers are measured by looking at *how much difference* the cause makes.[14] If studying philosophy in your 20s makes a slight difference in your chance of living until 50, its causal power is slight. But making little difference is perfectly consistent with the conditional probability's being high.

We now can address the question of whether fitness is a propensity. Again, for simplicity, let's focus on viability fitnesses. As noted earlier, the viability fitness of a trait is a conditional probability that has the form *Pr*(*O* survives to adulthood|*O* has trait *T*). The question is whether this conditional probability describes *T*s causal power to promote survival.

It does not. One reason comes the fact that the value of the conditional probability *Pr*(*O* survives to adulthood|*O* has trait *T*) tells you nothing about how much difference having trait *T* makes in the survival of an individual. It is difference making that represents the trait's causal power to influence survivorship, not the bare value of this single conditional probability. A second reason comes from Proposition (1), which friends of the propensity interpretation of fitness endorse. Two coextensive traits must have identical fitnesses, even if one promotes survival while the other is neutral or deleterious. That is the point of the selection toy I described in Sober (1984, p. 99). In this toy, there are balls of different sizes; the toy is divided into levels by barriers that contain holes; the holes are largest at the top level, somewhat smaller one level down, and so on. If all the balls are at the top and the toy is shaken, the balls get sorted, with the smallest balls finding their way to the bottom and the larger balls remaining closer to the top. Suppose a ball's "fitness" is its chance of getting to the bottom. Since all and only the red balls are the smallest, the two traits have identical fitnesses, though it is size, not color, that causes "success" in this process. For a biological example, suppose that having red blood and having hemoglobin in the blood are coextensive traits in a population. This means the traits have identical finesses, even though they play different causal roles. Having hemoglobin in your blood promotes survival (because it transports oxygen to tissue), whereas having red blood does nothing to help or hinder survival.[15]

The two arguments against regarding the viability fitness of a trait as a propensity the trait has to promote survival also apply to fertility fitness. For example, in cases where fertility fitness is adequately represented by expected number of offspring, that fitness also obeys the constraint that coextensive traits must have the same fitness value, and the expected number of offspring, conditional on having trait *T* says nothing about how having trait *T* causally influences reproductive success.

The fact remains that fitness is a probabilistic property, one that is not identical with actual survival or reproductive success. But it is not a propensity; by this I mean that the fitness of a trait should not be identified with the trait's causal power to promote survival or reproductive success.

Mills and Beatty (1979, p. 276) say that "intuitively, the fitness of a type (genotype or phenotype) reflects the contribution of a particular gene or trait to the expected descendant contribution…;" here they seem to endorse the idea that trait fitnesses are propensities. But earlier in their paper (p. 272), they claim that the fitness of a trait is not a propensity because only concrete particulars have propensities. I agree with the claim but not with the reason. Properties (types) have causal powers. For example, smoking cigarettes causally promotes lung cancer; individuals place themselves at greater risk by having this property. Talking about the causal power of a trait is elliptical for a claim about what an individual's having the trait would causally promote. Although traits have propensities (causal powers), the fitness of a trait is not the propensity the trait has to promote survival and reproductive success.

4. Where's the causality in natural selection?

What, then, becomes of the idea that natural selection causes evolution? If selection occurs precisely when there is variation in fitness, doesn't the failure of the propensity interpretation of fitness mean that selection is not a cause of evolution? This does not follow. We so far have asked whether the fitness of *a trait* represents the trait's causal power to promote the survival and reproduction *of an organism*. This does not address the question of whether *variation in fitness among traits* causally promotes the evolution of a *population*. Notice that the two questions involve different *relata*. The fact that the answer is *no* to the first does not entail that the answer is *no* to the second.

One causal aspect of selection can be seen by considering the distinction between selection-of and selection-for (Sober, 1984). In the selection toy just described, there is selection for being small, not selection for being red, even though the two traits are coextensive. There is selection for being small because a ball's being small causally promotes its getting to the bottom. There is no selection for being red, since being red plays no such causal role. When there is selection for one trait and selection against another, the traits *make a causal difference* in survival and reproductive success. This isn't something that is captured by the absolute fitness value of either trait. In the expression "selection for trait *T*," the predicate "*T*" occurs nonextensionally. This is no surprise, since the predicate also occurs nonextensionally in "*a*s having trait *T* causes…"

Let us now set aside the concept of selection-for and consider the concept of variation in fitness. Does fitness variation in a population causally promote change in trait frequencies? Again, the

[12] For two attempts to shield the propensity interpretation from Humphreys' objection, see Gillies (2000), Miller (2002).

[13] Humphreys (1985) also defends a more radical thesis; he says that "propensities must, it is commonly held, have the properties prescribed by probability theory. My contention is that they do not and, that rather than this being construed as a problem for propensities, it is to be taken as a reason for rejecting the current theory of probability as the correct theory of chance (p. 557)." Humphreys wants a new theory of chance, distinct from the current theory of probability. I take no stand on this thesis, but merely focus on his thesis that "propensities cannot be probabilities," which is the title of his paper. At the end of his paper Humphrey decries the practice of thinking of probability in terms of a set of abstract axioms (e.g., Kolmogorov's) and then asking how these axioms can be interpreted (p. 569). My reading of his counterexample to the propensity interpretation does not depend on buying the Kolmogorov axioms.

[14] The difference-making idea needs to be spelled out carefully. The strength of *C*'s propensity to bring about *E* (where both are dichotomous) should not be defined as $[Pr(E|C) - Pr(E|\text{not}C)]$, since that would equate strength of propensity with degree of correlation. The idea of holding fixed other causal factors, manipulating the one of interest, and recording the change in the probability of the effect is a natural suggestion. A further question is whether arithmetic differences in probabilities, or their ratios, should be used (Fitelson & Hitchcock, 2011). In many contexts there may be no clear answer; in the case of evolution by natural selection, there often is, as we shall see.

[15] It is curious that Sober (1984) endorses the propensity interpretation of fitness while at the same time highlighting the extensionality of fitness, as revealed by proposition (1).

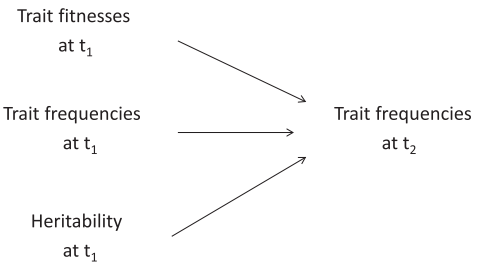

Fig. 2. A causal model in which variation in fitness at one time is a causal variable and trait frequencies at a later time is the effect variable.

selection toy is instructive. If it is size, not color, that causes different balls to descend to different levels in the toy, surely variation in *color* does not cause the balls to distribute themselves unequally. It is variation in *size* that is doing the causal work. Does this mean that it isn't variation in fitness that causes evolution—that the only concept that gets at causes is selection-for?

Not at all. Although the claim that A is fitter than B does not entail that having trait A (rather than B) causes organisms to be more successful at surviving and reproducing, it remains true that variation in fitness causes evolution. When mathematical models represent the "strength of selection" and distinguish that quantity from "response to selection," they are saying something causal, and they do so by adverting to variation in fitness. They do not use the concept of selection-for. Here's a simple example. Consider an asexual population in which there are two traits A and B with frequencies p and q and fitnesses $w(A)$ and $w(B)$, where offspring always resemble their parents. The relevant equation for the expected change in trait frequency in the next generation is (Crow & Kimura 1970, p. 179):

$$\Delta p = \frac{spq}{\bar{w}}.$$

The selection coefficient s represents the fitness difference $[w(A) - w(B)]$ and $\bar{w}$ is the average fitness of individuals in the parental generation (see footnote 14). Apparently, the strength of selection is a cause and the response to selection is an (expected) effect. A similar point applies to the breeder's equation in quantitative genetics (Falconer & MacKay, 1996):

$$\text{response to selection} = \text{heritability} \times \text{strength of selection}.$$

Strength of selection is calibrated by a measure of how much variation in fitness there is in the population; the response should be understood as an expectation.[16]

It is natural to represent the idea that variation in fitness is a cause of trait evolution in the kind of causal model depicted in Fig. 2. The three causes represented there are independent of each other. Each can be manipulated in such a way that the change effected in a causal variable is expected to be associated with a change in state of the effect variable (Woodward, 2003).

If trait A is fitter than trait B in a population, it does not follow that there is selection for A and against B. What does follow is that there is *some property or other* that A individuals have more frequently than B individuals do (call it P) and there is selection for P. If the red balls in the toy are more prone to descend than the green ones are, there must be some property that distinguishes them that is doing the causal work; that property need not be color. Where there is variation in fitness, there is also selection-for.

The proposition that there is selection for P and against Q is transparently causal; the claim that A and B vary in fitness is not. This may suggest that variation in fitness does not cause evolution, but this suggestion is mistaken. The claim that A and B differ in fitness *entails* that there is some property that A individuals have more frequently than B individuals and there is selection for that property. A claim of fitness difference must have causal content if it entails a proposition that is transparently causal.

5. Concluding comments

Although the propensity interpretation of fitness distinguishes the fitnesses of token organisms and the fitnesses of traits, where these two concepts are linked by Proposition (1), the motivation for the interpretation begins by thinking about the fitnesses of token organisms. The starting thought is that a token organism's fitness represents the organism's ability to survive and reproduce in its environment. That fitness gets represented by a conditional probability (if we restrict ourselves to considering the organism's viability). So the probability represents an ability—the causal power the organism has to survive to adulthood. But the causal *oomph* of the fitness concept starts to diminish when we think about the fitnesses of traits that numerous individuals share. The fitness of trait T is the average ability that individuals with trait T have of reaching adulthood. We have seen that this average ability of the individuals that have trait T does not represent the causal contribution that having trait T makes to surviving and reproducing. So much for the idea that the fitness of a trait represents its propensity to causally promote survival and reproduction.

There is causation in natural selection, but you have to look for it in the right places.[17] The fitness value of a trait does not represent the causal role the trait plays in affecting survival and reproductive success. But the concept of selection-for is causal. And it's also true that fitness variation[18] causally promotes evolution (Millstein, 2006; Sober, 2011a).

What about the concept of propensity? Though the fitness of a trait is not a propensity the trait has to promote survival and reproduction, the fitness variation in a population is a propensity the population has to evolve in a certain direction; the magnitude of that variation represents the causal push that selection gives to trait frequencies. [19] The propensity idea was on the right track; it just needs to be frame-shifted.

Acknowledgments

This paper is dedicated to the memory of James F. Crow, mentor and friend. My thanks to Hayley Clatterbuck, Maria Carla Galavotti, Casey Hart, Casey Helgeson, Matt Kopec, Joshua Mund, Steve Orzack, Grant Ramsey, Elena Spitzer, Mike Titelbaum, Olav Vassend, Joel Velasco, and the anonymous referees of this journal for helpful comments.

[16] In a similar vein, Fisher's (1999) fundamental theorem of natural selection says that the rate of change in fitness in a population due to natural selection equals the additive genetic variance in fitness (Okasha, 2008). Here variation in fitness is a causal propensity that the population has—not just to change trait frequencies, but to improve fitness. Fisher's theorem does not apply to frequency dependent fitnesses and in this sense it is less general than the breeder's equation. But its characterization of the causal power of variance in fitness is worth noting.

[17] The "statistical interpretation" of selection (Lewens, 2004; Walsh, 2000, 2007; Walsh, Lewens, & Ariew, 2002) rejects the causal interpretation; see Stephens (2004), Shapiro & Sober (2007) for replies to some of these papers. The present paper is not a response to this denial of causation; rather, it is aimed at those who already accept the idea that selection causes evolution.

[18] I use the term "variation" here without specifying how this is to be measured mathematically. I leave it open that different problems may require different measures.

[19] This propensity often cannot be interpreted as a probability. Variance in fertilities often takes values that are greater than unity. This is another example of what Humphreys (1985) was describing—a propensity that is not a probability.

References

Beatty, J., & Finsen, S. (1989). Rethinking the propensity interpretation—A peek inside pandora's box. In M. Ruse (Ed.), *What the philosophy of biology is* (pp. 17–30). Dordrecht: Kluwer.

Benacerraf, P. (1973). Mathematical truth. *Journal of Philosophy, 70*, 661–679.

Brandon, R. (1978). Adaptation and evolutionary theory. *Studies in the History and Philosophy of Science, 9*, 181–206.

Brandon, R. (1990). *Adaptation and environment*. Princeton: Princeton University Press.

Crow, J., & Kimura, M. (1970). *An introduction to population genetics theory*. Minneapolis: Burgess.

Drouet, I., & Merlin, F. (in press). The propensity interpretation of fitness and the propensity interpretation of probability. *Erkenntnis*.

Falconer, D., & MacKay, T. (1996). *Introduction to quantitative genetics*. Edinburgh: Pearson.

Fetzer, J. (1982). Probabilistic explanations. *Proceedings of the philosophy of science association biennial meeting, 2*, 194–207.

Fisher, R. A. (1999). *The genetical theory of natural selection* (Variorum edition 1999). Clarendon Press.

Fitelson, B., & Hitchcock, C. (2011). Probabilistic measures of causal strength. In P. Illari, F. Russo, & J. Williamson (Eds.), *Causality in the sciences*. Oxford: Oxford University Press.

Galavotti, M. C. (2005). *A philosophical introduction to probability*. Stanford: CSLI Publications.

Giere, R. (1973). Objective single-case probabilities and the foundations of statistics. In P. Suppes, L. Henkin, A. Joja, & G. Moisil (Eds.), *Logic, methodology and philosophy of science IV* (pp. 467–483). Amsterdam: North Holland Publishers.

Gillespie, J. (1977). Natural selection for variances in offspring numbers—A new evolutionary principle. *American Naturalist, 111*, 1010–1014.

Gillies, D. (2000). *Philosophical theories of probability*. London: Routledge.

Hájek, A. (2011). Interpretations of probability. In E. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. URL: <http://plato.stanford.edu/entries/probability-interpret/>.

Humphreys, P. (1985). Why propensities cannot be probabilities. *Philosophical Review, 94*, 557–570.

Lewens, T. (2004). *Organisms and artifacts*. Cambridge: MIT Press.

Miller, D. (1994). *Critical rationalism—A restatement and defence*. Open Court.

Miller, D. (2002). Propensities may satisfy Bayes's theorem. In R. Swinburne (Ed.), *Bayes's theorem* (pp. 111–116). Oxford University Press.

Mills, S., & Beatty, J. (1979). The propensity interpretation of fitness. *Philosophy of Science, 46*, 263–286.

Millstein, R. (2006). Natural selection as a population-level causal process. *British Journal for the Philosophy of Science, 57*, 627–653.

Okasha, S. (2008). Fisher's 'Fundamental Theorem' of natural selection: A philosophical analysis. *British Journal for the Philosophy of Science, 59*, 319–351.

Parker, G. (1978). Search for mates. In J. Krebs & N. Davies (Eds.), *Behavioral ecology—An evolutionary approach*. Blackwell.

Peirce, C. S. (1910). "Notes on the doctrine of chances", In Hartshorne, C., & Weiss, P. (Eds.), Collected papers of Charles Sanders Peirce, 1931 (vol. 2, pp. 405–414). Cambridge: Harvard University Press.

Pence, C., & Ramsey, G. (2013). A new foundation for the propensity interpretation of fitness. *British Journal for the Philosophy of Science*.

Popper, K. (1959). The propensity interpretation of probability. *British Journal for the Philosophy of Science, 10*, 25–42.

Popper, K. (1990). *A world of propensities*. London: Thoemmes.

Ramsey, G. (2006). Block fitness. *Studies in the History and Philosophy of the Biological and Biomedical Sciences, 37*, 484–498.

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* (pp. 235–264). University of Pittsburgh Press.

Sober, E. (1981). Evolutionary theory and the ontological status of properties. *Philosophical Studies, 40*, 147–176.

Sober, E. (1984). *The nature of selection*. Chicago: University of Chicago Press.

Sober, E. (1993). *Philosophy of biology*. Boulder: Westview Press.

Sober, E. (2001). The two faces of fitness. In R. Singh, D. Paul, C. Krimbas, & J. Beatty (Eds.), *Thinking about evolution: Historical, philosophical, and political perspectives* (pp. 309–321). Cambridge University Press.

Sober, E. (2011a). *A* priori causal models of natural selection. *Australasian Journal of Philosophy, 89*, 571–589.

Sober, E. (2011b). *Did Darwin write the origin backwards?* Amherst, New York: Prometheus Books.

Stephens, C. (2004). Selection, drift, and the 'Forces' of evolution. *Philosophy of Science, 71*, 550–570.

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. (2007). The pomp of superfluous causes: The interpretation of evolutionary theory. *Philosophy of Science, 74*, 281–303.

Walsh, D., Lewens, T., & Ariew, A. (2002). The trials of life—Natural selection and random drift. *Philosophy of Science, 69*, 452–473.

Woodward, J. (2003). *Making things happen*. Oxford: Oxford University Press.