Probability in Biology: The Case of Fitness

Roberta L. Millstein
Department of Philosophy
University of California, Davis
Davis, CA 95616
RLMillstein@UCDavis.edu

1. Introduction

The biological sciences make extensive use of probabilities, whether evolutionary biology, ecology, molecular biology, developmental biology, genetics, physiology, paleobiology, medicine, neurobiology, etc., and whether as parts of formal theoretical models or as parts of less formal mechanistic models. Organisms just don't seem to behave in determinate and fully regular ways. In principle, then, the discussion of probability in biology is a massive topic.

For example, evolution is probabilistic in the production of new variations (e.g., mutation; see Merlin 2010), in the microevolutionary perpetuation of those variations through natural selection and random drift, and even at the level of macroevolution; associated with these probabilities are a number of colloquial meanings of chance, such as "chance as ignorance," "chance as coincidence," or "chance as contingency" (see Millstein 2011 for an overview of the concept of chance in evolution). Models in ecology are also thoroughly probabilistic, generally eschewing most lower-level details for equations containing only a few simple parameters (Colyvan 2005). In neurobiology, many phenomena are probabilistic, such as whether an Na^+ channel will open or a synapse will be potentiated; often these probabilities are low ones (Craver 2007). And any

student of medicine is aware of the ubiquity of probabilities, particularly where uncertainty is involved (Djulbegovic, Hozo, and Greenland 2011).

However, most of the discussion of probability in the philosophy of biology has focused on probability in evolution, and much of that has centered on the concept of "fitness" (followed closely, perhaps, by discussions of "random drift"). In part, that is a reflection of the philosophy of biology's focus on evolution more generally, a focus that is now shifting to embrace the other areas of biology. Thus, it is somewhat apologetically that I find myself perpetuating rather than bucking the older trend.[1] My reasoning is this: first, by discussing the probabilistic nature of fitness, I will be able to discuss a wide ranging, established literature; and second, because the literature is so well established, it can serve as a guideline for where discussions of probability in biology can go wrong and where they can go right. Indeed, my goal in this essay is to show how discussions of fitness in the philosophy of biology have gone profoundly off track, and to try to gently nudge them back on track, pointing out important insights that were missed along the way.

I begin with a discussion of "fitness" in the work of Charles Darwin, which will also serve to introduce several key themes which are echoed in the contemporary discussion. I then provide an overview of the account of fitness that has garnered the lion's share of discussions about fitness, the propensity interpretation of fitness, followed by some critiques that have been offered of it. Many of these critiques have focused on the mathematical aspects of fitness rather than

---

[1] A second apology: although I attempt to survey a fair bit of the literature on fitness, the topic alone is so huge that (for reasons of space) there will be papers that I do not cite and concerns that I do not discuss. I have instead sought to focus on what I take to be the most central and interesting issues, though no doubt some of those have been inadvertently left out as well.

fitness as a propensity per se; I proceed to show how re-focusing on several aspects of the propensity interpretation of probability more generally can help to address the concerns that have been raised. I conclude with some general lessons for thinking about probability in biology.

2. Background: Darwin and Fitness

Perhaps surprisingly, the precise term "fitness" is not used in anything like its current evolutionary context in any of the six editions of Darwin's *Origin of Species*. Of course, the phrase "survival of the fittest" *is* well known;[2] Darwin used the phrase, due to Herbert Spencer, at the urging of Alfred Russel Wallace. The phrase made its first appearance in Darwin's *The Variation of Animals and Plants under Domestication*, and the fifth edition of the *Origin* was the first edition to contain it. However, it is instructive to see how Darwin uses very similar terms, such as "fit," "fitting," or "fitted," from the very first edition of the *Origin*. His uses are far too numerous to list here, but here are some typical ones:

> Let it be borne in mind how infinitely complex and *close-fitting* are the mutual relations of all organic beings to each other and to their physical conditions of life (Darwin 1859: 80; emphasis added).

> Or, again, the wolves inhabiting a mountainous district, and those frequenting the lowlands, would naturally be forced to hunt different prey; and from the continued preservation of the individuals best *fitted* for the two sites, two varieties might slowly be

---

[2] Well known, and somewhat notorious, as the phrase has caused no end of trouble (e.g., charges, albeit false charges, that the theory of natural selection is circular).

formed (Darwin 1859: 91; emphasis added).

I look at all the species of the same genus as having as certainly descended from the same progenitor, as have the two sexes of any one of the species. Consequently, whatever part of the structure of the common progenitor, or of its early descendants, became variable; variations of this part would, it is highly probable, be taken advantage of by natural and sexual selection, in order to *fit* the several species to their several places in the economy of nature, and likewise to *fit* the two sexes of the same species to each other, or to *fit* the males and females to different habits of life, or the males to struggle with other males for the possession of the females (Darwin 1859: 157-8; emphasis added).

One important thing to notice about these quotes is that being fit is a *relational* term. That is, organisms can be fitted to other organisms or to their "physical conditions of life" (external environment). More specifically, they can be fitted to their places in the "economy of nature" (i.e., the roles they play in an ecosystem), to members of the opposite sex in the same species, to their "habits of life," or for combat with members of the same sex for access to mates. In this way, being "fit" in an evolutionary context differs from the colloquial usage of the term; an organism is only fit relative to some aspect of the environment or to other organisms, not fit in an absolute sense. This can be seen most clearly in the wolf quote. Neither variety of wolf is absolutely fitter than the other; rather, one is fitter for the lowlands and one is fitter for the mountainous region. Contemporary views of fitness may differ on many points, but not on this one; fitness is always relative to the environment and to other organisms.[3] (I suspect that Darwin

---

[3] We will see below that fitness can be construed as relational in another sense as well, in *comparing* the fitness of one organism (or type) to another, as in the claim "*X* is fitter than *Y*." Consider "weight" as an

was referring to organisms being fitted to members of *other* species, as mistletoe is to certain trees, birds, and insects, one of Darwin's examples. Later, we will have reason to consider whether fitness ought to be relative to organisms of the same species as well).

The discussion of Darwin's views on fitness thus far makes it sound as though fitness (and by implication, natural selection) is a solely a matter of organisms' abilities to survive without consideration of their abilities to reproduce, and indeed, much of the *Origin* reads that way. In particular, Darwin devotes an entire chapter (Chapter 3) to discussing what he called the "struggle for existence"; the chapter is mostly about the struggle to survive. However, reproductive ability does get brief and occasional mention; two oft-quoted passages are "I should premise that I use the term Struggle for Existence in a large and metaphorical sense, including dependence of one being on another, and including (which is more important) not only the life of the individual, but *success in leaving progeny*" (1859: 62; emphasis added) and "can we doubt (remembering that many more individuals are born than can possibly survive) that individuals having any advantage, however slight, over others, would have the best chance of surviving and of *procreating their kind*?" (1859: 81; emphasis added). In the *Descent of Man*, Darwin is a bit more clear about including reproduction as part of his explanation of how to distinguish natural selection from sexual selection (see Millstein 2012 for discussion), although even there, he slips

analogy. An object has the weight that it does in virtue of being in a particular gravitational field. That is one sense of "relational," where "relational" means *context dependent*. However, one can also speak of an object as being "heavy," which implies that it weighs more than another object (in a specified gravitational field). This seems to be a slightly different sense of "relational": *comparative*. In discussing the relation of fitness to the environment and other organisms, I mean to invoke only the context dependent sense. Thanks to Marcel Weber for clarification on this point.

occasionally.[4] In any case, in contemporary evolutionary biology, reproductive ability is

generally considered to be part of fitness,[5] although sometimes *viability* and *fecundity* are

considered separately.

Finally, notice that I have repeatedly referred to fitness in terms of survival *ability* and

reproductive *ability.* This is because Darwin says, repeatedly, that individuals having any

advantage over others would have the *best chance* of surviving and of procreating their kind (see,

e.g., quote in the preceding paragraph). That is, the fitter organisms – the organisms with

"structures" or other characteristics that provide advantages[6] over other organisms – may not in

fact be the organisms who have the greater success in surviving and/or reproducing. Thus, from

the very beginning, fitness (and thus natural selection) was seen as a *chancy* affair (Beatty 1984,

Richardson and Burian 1992, Millstein 2002), paving the way for a contemporary understanding

of fitness as probabilistic.

Although a full history of the term "fitness" is beyond the scope of this essay (but see, e.g.,

Gayon 1998 and Jackson manuscript), it is important to understand that when evolutionary

theory became mathematized in the early 20th century in the form of population genetics (see

Millstein and Skipper 2007 for discussion), fitness became a parameter used in equations. In

other words, it became important to *quantify* organisms' fitnesses, to be able to say how much

fitter some organisms were than others and to be able to use different fitness values to predict the

---

[4] For example: "Sexual selection acts in a less rigorous manner than natural selection. The latter produces its effects by the life or death at all ages of the more or less successful individuals" (1871 vol. I: 278-9).
[5] See Gayon (1998) for a discussion of how contemporary evolutionary biology came to emphasize the reproductive aspect of fitness and the role that eugenics played in that discussion.
[6] On Darwin's view, these advantages were usually seen as "slight" or "small"; over time, natural selection would accumulate those small fitness differences into adaptations.

frequencies of different types in the next generation. This, too, will be important for understanding contemporary debates over fitness.

3. The Propensity Interpretation of Fitness

Contemporary philosophical discussions of fitness generally begin with Susan Mills and John Beatty's classic 1979 article, "The Propensity Interpretation of Fitness," though proper credit must also be given to the Robert Brandon's (1978) independently and simultaneously developed propensity account of fitness (or, as he terms it, "adaptedness"). By the time of Mills and Beatty's article, most biologists were defining fitness in terms of an organism's *actual* survival and reproductive success – in other words, the actual number of offspring produced by an organism. Let us call this the *actualist definition of fitness*. Claims that evolutionary theory is untestable had been made and refuted, but the actualist definition of fitness caused evolutionary explanations invoking fitness to be circular, Mills and Beatty argued. If we seek to explain why type *A* is leaving a higher number of offspring than type *B*, the purported explanation "because *A* is fitter than *B*" is circular if "*A* is fitter than *B*" amounts only to the claim that "*A* left more offspring than *B*," as it does on the actualist definition of fitness. Mills and Beatty sought to solve this problem of *explanatory circularity* with the *propensity interpretation of fitness*.[7]

However, explanatory circularity was not the only problem that Mills and Beatty found with the actualist definition; they also noted that there was a mismatch between the actualist definition – that is, biologists' stated definition of fitness – and biologists' *usage* of the term fitness. Let us

---

[7] Mills and Beatty's main goal was *not* to try to address the supposed circularity of the phrase "the survival of the fittest," although they thought that it could; they wisely note that "[t]his catch-phrase is not an important feature of evolutionary theory" (Mills and Beatty 1979: 270, n. 5).

call this the *mismatch problem*. They used two examples to illustrate the mismatch problem. The first example draws on what has become an oft-cited scenario from Scriven (1959). A pair of identical twins are standing together in a forest; one is struck by lightning before reproducing while the other emerges unscathed and is later able to reproduce successfully.[8] The actualist definition of fitness would have us say that the second twin is far fitter than the first, since the actual reproductive success of the second twin is far greater than that of the first twin (whose reproductive success is zero). Yet, Mills and Beatty urge, this is counter-intuitive. If the twins are phenotypically and genetically identical (which I think is the supposition of the example), how can it make sense to say that one is fitter than the other? No biologist would actually use the term fitness in that way. The second example is also hypothetical but is somewhat more realistic:

> Imagine two butterflies of the same species, which are phenotypically identical except that one (*C*) has color markings which camouflage it from its species' chief predator, while the second (*N*) does not have such markings and is hence more conspicuous. If *N* nevertheless happens to leave more offspring than *C*, we are committed on the definition of fitness under consideration to conclude that (1) both butterflies had the same degree of fitness before reaching maturity (i.e., zero fitness) and (2) in the end, *N* is fitter, since it left more offspring than *C* (Mills and Beatty 1979: 268, n. 2).

And surely, it cannot be that the more conspicuous butterfly is fitter than the camouflaged butterfly! No biologist would use the term fitness in that way. The actualist definition of fitness

---

[8] This would not be an instance of selection – or drift, for that matter – since there are no differences between the twins, and both selection and drift require variation.

is inadequate; it is mismatched with biologists' usage of the term.

Mills and Beatty's solution to the explanatory circularity problem and the mismatch problem is quite elegant and hearkens back to Darwin's way of thinking. The essence of their account is that fitness refers not to an organism's *actual* survival and reproductive success, but rather to its *ability* (i.e., disposition, capability, or tendency) to survive and reproduce. Thus, the twins were equal in fitness, whereas the camouflaged butterfly was fitter than the conspicuous butterfly. Actual reproductive success may be an indicator of an organism's ability, but it is a defeasible indicator; it can happen that the less fit butterfly out-reproduces the fitter butterfly. As Darwin seems to suggest, the fitter organism has the *best chance* of out-surviving and/or out-reproducing the less fit, but it may not do so. In other words, an organism has a *probabilistic* ability to survive and reproduce, i.e., it has a *propensity* to survive and reproduce.[9]

What does this propensity consist in? When we speak of other complex dispositional properties, such as solubility of salt in water (i.e., the propensity of salt to dissolve in water), we refer the *physical properties* of the salt (its ionic crystalline character) in the presence of the appropriate triggering conditions and in the absence of any countervailing conditions. Similarly, the fitness of an organism consists in its having physical traits that lend themselves to survival and/or reproduction in a particular environment; again, the camouflaged butterfly in a similarly colored environment can serve as an example. For both the salt and the butterfly, the presence of the appropriate physical properties of the entity, together with the presence of certain triggering conditions in the absence of countervailing conditions, causes (probabilistically, in the case of

---

[9] The probabilistic nature of natural selection (and thus fitness) is distinct from probabilistic random drift; see Beatty (1984) and Millstein (2002) for discussion.

fitness) the manifestation of the specified behavior.

Thus, not only does the propensity interpretation solve the mismatch problem by proposing a view of fitness that is consistent with biologists' usage of the term, but it also addresses the explanatory circularity problem. With the propensity interpretation, if we seek to explain why type *A* is leaving a higher number of offspring than type *B*, the explanation "because *A* is fitter than *B*" means that *A* has a greater propensity than *B* to survive and reproduce in the given environment, which means that the physical properties of *A* in its environment are what cause it to tend to have greater reproductive success than *B* (with its physical properties). The relative physical abilities can be determined by engineering optimality models or other examinations of the physical properties of the organism in its environment, and then be confirmed by measurements of actual descendant contributions (Mills and Beatty 1979: 277; see also Brandon and Beatty 1984 and Brandon 1990 on this point). Also, Mills and Beatty's account of fitness incorporates the features of Darwin's discussed above: fitness is relative to the environment, fitness involves both survival and reproduction, and fitness is a chancy ability.

The propensity interpretation of fitness quickly gained a number of prominent adherents in addition to Mills, Beatty, and Brandon, such as Richard Burian (1983), Elliott Sober (1984), Philip Kitcher (1987), Ernst Mayr (1988), and Robert Richardson and Burian (1992). Of course, there were critics, too (Rosenberg 1985, Rosenberg and Williams 1986, Hodge 1987). Anyone familiar with the philosophy of probability might think that the lion's share of the criticisms came from proponents of other interpretations of probability, e.g., frequentist interpretations or epistemic interpretations. In fact, although there has been some discussion of alternate

interpretations of probability for fitness and analysis of the propensity interpretation's appropriateness for the concept of fitness (e.g., Waters 1986; Richardson and Burian 1992; Bouchard and Rosenberg 2004; Abrams 2006, 2007), there has been surprisingly little.[10] Indeed, I would say that the fact that there have been so few publications on interpretations of probability as compared to publications on the topic of fitness more generally is one indication that the discussion has gone in some strange directions.

One might also expect that there would be some discussion of whether Mills and Beatty's proposed propensities were "true" (indeterministic) propensities, akin to the propensities involved in the radioactive decay of atomic particles. Mills and Beatty's view seems to have been that they are (1979: 273). Indeed, there has been a fair bit of discussion of whether evolution is indeterministic, with perhaps more authors arguing that it is than arguing that it isn't. My own view on this is that one can make sense of both deterministic and indeterministic propensities (Millstein 2003a) and that there is reason to be agnostic on the question of evolutionary indeterminism, at least for the time being (Millstein 2003b; this piece contains a review of the literature on evolutionary indeterminism up to that point).

4. The Propensity Interpretation of Fitness Takes a Mathematical Turn

Perhaps the view of fitness as a propensity would have rested there were it not for mathematical

---

[10] Some of the options can be ruled out quickly. An epistemic interpretation of probability would not be explanatory in this context, i.e., it would not explain why a population changed in frequency from one generation to the next. And a frequentist view that interpreted the probabilities in terms of *actual* frequencies would not solve the explanatory circularity problem. See Brandon (1990) and Richardson and Burian (1992) for a discussion of the problems with using a limit or long-run relative frequency interpretation for fitness.

population genetics models referred to at the end of Section 2. Again, it is generally not enough to say that *A* is fitter than *B*; one must also say *how much* fitter *A* is than *B*: 50% fitter? Twice as fit? etc. Note also that whether an organism reproduces or fails to reproduce is often not what is most relevant for natural selection. Reproduction is not an all or nothing affair; it matters *how many* offspring an organism has. Consider, for example, an organism who has zero offspring, an organism who has one offspring, and an organism who has twelve offspring. The effect on the distributions of types in the next generation of having one offspring is more similar to having zero offspring than it is to having twelve offspring. Furthermore, as Mills and Beatty point out, an organism can have different propensities to have different numbers of offspring. Perhaps one organism has a 25% chance of having zero offspring, a 50% chance of having one offspring, and a 25% chance of having twelve offspring (again, based on its physical traits in the particular environment). How does its fitness compare to a second organism who has a 25% chance of having zero offspring, a 25% chance of having one offspring, a 25% chance of having 8 offspring, and a 25% chance of having twelve offspring? What changes in gene frequencies would we expect in the next generation based on these propensity values?

Mills and Beatty, despairing of the possibility of comparing one such distribution of reproductive propensities to another, suggest that the fitness value for an *individual organism* in its environment ($fitness_1$) be an organism's expected number of offspring in that environment, where the expected number is the weighted sum of the possible offspring contributions. In the previous paragraph, the $fitness_1$ of the first organism is $.25*0 + .5*1 + .25*12 = 3.5$ whereas the $fitness_1$ of the second organism is $.25*0 + .25*1 + .25*8 + .25*12 = 5.25$. In other words, the second organism is fitter than the first because its expected number of offspring is greater: 5.25

as compared to 3.5. Mills and Beatty point out that biologists also often want to talk about the fitness of *types* (genotypes or phenotypes)*,* where the fitness "reflects the contribution of a particular gene or trait to the expected descendent contribution (i.e., the fitness$_1$) of possessors of the gene or trait" (1979: 276). However, organisms possessing the gene or the trait in question will vary in their other genes or traits, and these genes or traits will affect the organism's fitness; thus, they define fitness$_2$, the fitness of types, as the *average* expected number of offspring for a given genotype or phenotype (i.e., the average fitness$_1$). This allows us to "predict and explain the evolutionary fate of the genes or traits which correspond to the alternate types" (Mills and Beatty 1979: 276). Finally, since we are generally interested in predicting or explaining the success of one type as compared to others in the population, Mills and Beatty define *relative* fitness$_2$ of a type as the fitness$_2$ of that type divided by the type that has the highest fitness$_2$ in the population (so that the type with the highest fitness$_2$ will always have a relative fitness$_2$ of 1).[11]

4.1 Non-mathematical fitness or mathematical fitness?

It is hard to know how to reconcile fitness as a propensity, based on the heritable physical characteristics of an organism (i.e., the interpretation of fitness that addressed the explanatory circularity problem and the mismatch problem) with fitness$_1$, fitness$_2$, and relative fitness$_2$. In particular, there seems to have been a shift from the *causes* of an organism's survival and reproductive success (its physical traits in relation to its environment) to the expected *outcomes* (the expected survival and reproductive success) themselves. Are these different notions really part and parcel of the same propensity interpretation of fitness? Mills and Beatty do not offer us

---

[11] Note that this is "relational" in the sense of comparative, rather than context-dependent. (See note 3).

much guidance except to note that "the fitness values assigned to organisms [i.e., the expected reproductive success values] are not literally propensity values, since they do not range from 0 to 1"; instead, the fitness of an organism is "a complex of its various reproductive propensities" (1979: 275). Some philosophers of biology (e.g., Matthen and Ariew 2002, Pigliucci and Kaplan 2006), referring to the non-mathematical notion as "vernacular" fitness and the mathematical notion as "predictive" fitness, suggest that the two cannot be reconciled (and that the former is causal whereas the latter is not). Other philosophers of biology simply equate the propensity interpretation of fitness with the mathematical formulation (e.g., Abrams 2007), more or less ignoring the non-mathematical formulation, which avoids the issue altogether. Bouchard (2011) argues that a more general account of fitness will not focus on reproductive success. Brandon (1990) seems to differentiate what fitness *is* (how it is defined, what it is ontologically, which is the non-mathematical formulation) from how it is *measured* (the latter being consistent with Mills and Beatty's references to fitness$_1$, fitness$_2$, and relative fitness$_2$ as fitness *values*, i.e., the mathematical formulation, of which there may be many), which suggests that it is mistaken to think of the mathematical formulation of fitness as a definition of fitness at all (or as the propensity interpretation in particular). Sober (2001), however, argues that reflections on the mathematical formulation of fitness call the propensity interpretation into question (despite his earlier endorsement of it).

How should we sort through these widely varying reactions for dealing with what Sober (2001) has dubbed "the two faces of fitness"? And can the mathematical and non-mathematical aspects of fitness be reconciled after all? My own sympathies lie most closely with Brandon's response, and I think that some sort of reconciliation may in fact be possible; however, to address Sober's

rejection of Brandon's solution, we must first explore some of the mathematical complexities of using expected reproductive success as a proxy for distributions of propensities to produce different numbers of offspring.

4.2 Opening Pandora's Box

Ten years after propounding the propensity interpretation of fitness, Mills (now Finsen) and Beatty famously peeked inside Pandora's box and returned as critics, "to acknowledge and reframe some old problems, and to introduce some additional difficulties" (Beatty and Finsen 1989: 17). For a detailed discussion of these problems, I refer the reader to Beatty and Finsen's paper; here I seek only to summarize them. The first problem has to do with the time scale of fitness, a problem that Mills and Beatty (1979) addressed briefly in a pair of footnotes. Recall that fitness$_1$ is the expected number of offspring which an organism will leave in a given environment. But suppose organism x has a greater expected number of offspring than organism y, whereas y has a greater expected number of descendants than organism x in the subsequent generations. (Perhaps, for example, x cannot provide good offspring care for so many organisms, so that its offspring do not themselves have the ability to produce many offspring). Who is fitter, x or y? Mills and Beatty suggested that we differentiate "between long term fitness and short term fitness–or between first generation fitness, second generation fitness, . . ., n-generation fitness" (1979: 275, n. 10). Thus, it might turn out that while x is fitter in the short-term, y is fitter in the long-term. One might even consider as long a scale as "expected time to extinction" (Beatty and Finsen 1989: 18). Thus, perhaps we cannot say that there is one propensity interpretation of fitness, but rather, a (very large) family of fitness propensities.

A second set of problems discussed by Beatty and Finsen arises from the choice to use expected reproductive success – one value – to represent a distribution of different propensities to leave different numbers of offspring, even though there are other aspects of a distribution, such as its variance or its skew (asymmetry), that expected value does not reflect.[12] Their discussion derives from that of Gillespie (1974). Beatty and Finsen show that two genotypes can have the same expected reproductive success (fitness$_2$), but if one of the distributions has lower variance, that genotype will probably (i.e., is mathematically more likely to) be more reproductively successful than the other. Furthermore, if two genotypes have the same fitness$_2$ and their distributions have the same variance, but one has greater skew than the other, the one with greater skew will probably have greater reproductive success. Finally, in some cases, the greater skew of a distribution can compensate for larger variance, so that the genotype with greater skew but larger variance in its propensity distribution will probably have greater reproductive success. Thus, in all three of these cases, expected value yields the wrong prediction. In order to have a predictive calculation of future type frequencies for such cases, one must use the geometric mean rather than the arithmetic mean (i.e., the expected value). To make matters worse, whether we need to use arithmetic mean or geometric mean depends on whether there is variance within generations or between generations, respectively. But determining the distribution of propensities is hard enough; how will we determine whether the variance of the distribution manifests itself within generations or between generations, or even worse, whether organisms are using different strategies at different points in time? This is a problem, Beatty and Finsen assert, because "we may sometimes have no access to the sort of information we need in order to decide

---

[12] Again, I refer the reader to Beatty and Finsen's article for the technical details behind the discussion in this paragraph and the next, details that are a matter of mathematics rather than biology and which are uncontested as far as I know.

what statistic on the fitness distribution to employ in order to explain a particular evolutionary phenomenon" (1989: 28).

Beatty and Finsen suggest again that it seems as though there is a large family of propensities, rather than a single propensity. And then they conclude:

> It would be unfair to suggest that, lacking any generally agreed-upon definition of 'fitness', we therefore lack *any* understanding of evolution in terms of fitness differences and natural selection. On the other hand until we have an appropriate general definition of fitness, it is not altogether clear *how much* we understand about evolution in terms of fitness differences and natural selection (1989: 29; emphasis in original).

In short, Beatty and Finsen don't seem satisfied with a family of propensities of fitness; they suggest that greater understanding will be obtained only when we have a more general definition.

As with the original Mills and Beatty article, Beatty and Finsen's article has been quite influential, with a variety of conclusions being drawn, mostly arguing that the complications of variance and skew spell varying degrees of trouble for the propensity interpretation of fitness. Sober's response is particularly notable. Sober (2001), also drawing on the work of Gillespie (1974, 1977), argues that within-generation variance raises an additional problem, one that casts doubt on the propensity interpretation altogether. Sober points out that in order to make a predictive calculation of future frequencies when there is within-generation variance, one must include the *size of the population* in the calculation. But this is strange because the size of the

population is not causally affecting organisms' reproductive success (or so Sober claims – I will challenge this claim below). Thus, Sober concludes, "an organism's fitness is *not a propensity that it has*—at least not when fitness must reflect the existence of within-generation variance in offspring number" (2001: 320; emphasis added). Rather, Sober suggests, fitness is a "holistic" quantity that reflects both properties that affect an organism's reproductive success (the organism's traits in relation to its environment) and properties that do not (the size of the population).

At this point the reader may be forgiven for being confused. Short-term vs long-term reproductive success? Between-generation variance? Within-generation variance? Skew? What does any of this have to do with solving the explanatory circularity problem? What does this have to do with solving the mismatch problem? What do these problems of making mathematical predictions for future generations have to do with whether fitness is a physical property of organisms in their environments to survive and reproduce? In other words, what do these mathematical issues have to do with *the question of whether fitness ought to be construed as a propensity*? Are they even problems with propensities in particular? After all, *any* interpretation of probability is consistent with the use of probability distributions and their moments (a technical term referring to properties of a distribution like expected value, variance, and skew); under *any* interpretation of probability, the same problems with making predictions from probability distributions would still be present. It seems as though the debate has gotten off track. Perhaps a re-examination of the nature of propensities will bring us back on track.

5. Reflections on Propensities: Closing Pandora's Box?

Contemporary philosophers of probability often distinguish between two versions of the propensity interpretation, a long-run version and a single-case version:

> A long-run propensity theory is one in which propensities are associated with repeatable conditions, and are regarded as propensities to produce in a long series of repetitions of these conditions frequencies which are approximately equal to the probabilities. A single-case propensity theory is one in which propensities are regarded as propensities to produce a particular result on a specific occasion (Gillies 2000: 822; see also Gillies, this volume).

Mills and Beatty do not distinguish between these two possibilities, and it is not clear to me that it is important that they do so. Would it help with the question of whether fitness refers to success after one generation, two generations, or $n>2$ generations, for example? Although some have tried to argue for an exclusively short-run or an exclusively long-run view of fitness, it seems clear that they are different and that our views of fitness must account for both. It is also clear that the propensity for long-run reproductive success is not simply the summation of many identical short-run reproductive successes, since the traits that tend toward long-run reproductive success (e.g., adaptability to different conditions) may be different from those that tend toward single-generation reproductive success (e.g., high fecundity).

5.1 Propensity to *X*

What *is* important to note, however, is that either way, a propensity is a tendency to produce a particular sort of *behavior* (again, whether one single outcome or outcomes in the long-run). That is, a propensity is always a *propensity to exhibit a behavior X*. It makes no sense to talk of a propensity in the absence of specifying *X*. This may seem like a trivial point, but I think it will help us to understand the question of whether we should be talking about reproductive success after one generation or after some number *n* generations.

The ability of an organism to have *i* descendants after one generation describes one behavior. The ability of an organism to have *j* descendants after two generations describes another. The ability of an organism to have *k* descendants after some specified *n* generations describes a third behavior. These are, in fact, different propensities because they describe abilities for different behaviors. The different behaviors may, or may not, have different underlying physical traits that give rise to the ability.

Consider an analogy. Dara Torres has a propensity to swim fast. So does Janet Evans. Both were, at one point in time, U.S. swimmers over the age of 40 vying to compete in the 2012 Summer Olympics,[13] and both have won gold medals in previous Olympic competitions. Who is the faster swimmer? Well, Torres is a sprinter. In the shorter distances (e.g., 50m freestyle), she would likely prevail. But Evans, the long-distance swimmer, would likely beat Torres in the 1500m freestyle. So, should we be worried about whether "being a fast swimmer" is a propensity? No. There is no puzzle here. These are different abilities for different behaviors that in all likelihood have different underlying physical manifestations (e.g., perhaps more fast-

---

[13] Alas, neither succeeded in making the 2012 U.S. Olympic Swim Team.

twitch muscles for Torres, more slow-twitch muscles for Evans). In other words, the propensity

to succeed in the 50m freestyle is different than the propensity to succeed in the 1500m freestyle.

And the two propensities are equally legitimate.

So, which is the real propensity of fitness, the ability to have many descendants after one

generation or the ability to have many descendants after n generations? They are both equally

deserving of the name; they are different propensities and different abilities. As Beatty and

Finsen suggest, fitness should be understood as a family of propensities to survive and

reproduce. (Similarly, we can understand "Olympic swimmer" as a family of propensities to

swim fast). But at the end of their article, Beatty and Finsen bemoan the lack of a general view.

There is no need to bemoan the lack of a general view.[14] We should not expect one; indeed, one

would not be desirable if it led us to think the same set of physical traits gave rise to the ability to

have short-run reproductive success as gave rise to the ability to have long-run reproductive

success for a given genotype of a given species. Again, this is not likely, and certainly not

guaranteed always to be the case.

5.2 Propensities are Relational

Another important aspect of propensities is one that Karl Popper emphasized. On Popper's view,

propensities are not inherent in individual things, rather "they are *relational* properties of the

experimental arrangement" (1959: 37; emphasis added). As we have already discussed, it is well

agreed upon (since before Darwin and beyond) that fitness is relational. Darwin listed many

---

[14] Sober (2001) also argues that there is no problem with accepting both a long-term concept of fitness
and a short-term concept of fitness.

such relations, whereas contemporary accounts of fitness, such as that of Mills and Beatty, specify only that fitness refers to the ability of an organism to survive and reproduce in its environment. But what does the environment include? In particular, does it include the population of which the organism is a member?

Sober asserts that the size of the population plays no causal role except in "special cases" like density-dependent selection or frequency-dependent selection. He gives the example of four cows standing in the four corners of a large pasture; two of them have within-generation variance in their associated probability distributions and two don't. Sober states, "The cows are causally isolated from each other, but the fitnesses of the two strategies reflect population size" (2001: 32). Recall that this is supposed to be a problem for the propensity account of fitness – a serious enough problem that implies we ought to abandon the propensity interpretation of fitness altogether. However, there are two problems with this example and its underlying assumptions. First, cows alone cannot form a population; one needs bulls.[15] I don't mean to be snarky here, but this matters. Whether the cows are choosing their mates or whether the farmers are choosing their mates for them, who is mated with whom and how many mates are available will affect each cow's reproductive success (i.e., their fitnesses). Second, it is simply implausible that each cow stays in her own corner, never venturing to eat another cow's food, drink her water, or occupy her shelter, which again would affect the reproductive success of the other cows. In short, as I have argued elsewhere, what makes a population a population at all is that there are causal interactions – in particular survival and reproductive interactions – between the organisms (Millstein 2009, 2010); if the cows really are not interacting, then they are not part of the same

---

[15] Thanks to a certain Iowa boy for pointing out what ought to be obvious to us city folk, but isn't always.

population and not undergoing the same selection process.[16] It would then not make sense to compare their fitnesses. Or, to put the point in a more Darwinian vein, the struggle for existence, which is a result of various checks on the population size (due to limited resources, climate, prey, etc.), is part of the process of natural selection. Without it, organisms could reproduce without bounds. Sober's example, even if it were possible, is not an example of natural selection.[17]

In short, fitness is an organism's propensity to survive and reproduce in a given environment and in a given population (or, more precisely, is a family of propensities to do so) – and not just in "special cases" of density-dependence and frequency-dependence. In other words, an organism's fitness is relative to both its environment *and* to its population. (If that was always meant to be implicit, so be it.[18] The problem with leaving things implicit is that they can be forgotten). This suggestion is similar, although perhaps not identical, to that of Evelyn Fox Keller, who states, "The fitness of a particular female (or male) reproducing sexually always depends, first, on the availability, and second, on the fertility of males (or females) in the breeding group in which that organism finds itself," concluding that "[f]or sexually reproducing organisms, fitness is in general not an individual property but a composite of the entire interbreeding population" (1987: 393). I would generalize her reasoning to include all the ways that organisms affect each others' survival – both positively and negatively – in the struggle for existence.

---

[16] Ariew and Ernst (2009) and Walsh (2007) also overlook this point.
[17] See Lennox and Wilson (1994) for an argument for why "struggle for existence" is a necessary condition for natural selection, and why changes due *only* to differences in intrinsic rate of reproduction ought to be considered a different evolutionary process. Note that Ariew and Lewontin's (2004) claim that the "two faces of fitness" cannot be reconciled relies on conflating these two processes into one.
[18] See Mills and Beatty (1979: 271, n. 6).

If this is right, then it is not puzzling that a calculation for the future relative frequencies in a population should contain population size as a parameter. Features of the population, including its population size, are causal factors in natural selection.[19] Thus, contra Sober, the inclusion of population size in certain calculations is no reason to reject fitness as (a family of) propensities.

5.3 Determining the Value of a Distribution of Propensities

Even if the arguments of Section 5.2 are accepted, it might seem as though Pandora's box is still open; we still do not have a univocal way of calculating fitness values and thus no univocal way of saying whether one organism is fitter than another. Recall that the problem arises because for each organism, we can (in principle – this is hard to do in practice!) describe a distribution of propensities to have varying numbers of offspring. Using these probability distributions, we can easily calculate the expected reproductive success for each organism, the expected reproductive success for each type (by averaging all the organisms of that type), and the relative fitness of each type. However, in some cases, the expected value will give us a misleading result; A will have a higher expected reproductive success than B, but mathematically, B is likely to have a greater number of offspring than A. Again, as discussed above, the amount of variance or skew in a distribution affects the predicted number of descendants in future generations.

The basic problem, of course, is that we are trying to use one value to describe an entire distribution that contains many different characteristics – more technically, moments. Or as Beatty and Finsen put the point, "identifying fitness generally with any one statistic, or any

---

[19] This is also to be expected if one understands natural selection to be a form of discriminate sampling; see Beatty (1984) and Millstein (2002). Sampling processes, whether discriminate or indiscriminate, are always affected by population size.

particular function of statistics, is mistaken" (1989: 29). Each of those moments, whether

expected value, variance, skew, or any other moment of the distribution, captures some aspect of

the distribution while overlooking others. What we really want is to be able to compare one

entire probability distribution to another, but there does not seem to be a way to do this. Or is

there?

Recent papers by Sean Rice and colleagues (Rice 2008; Rice and Papadopoulos 2009; Rice,

Papadopoulos, and Harting 2011) suggest that there might be.[20] In their model, an individual's

fitness is the number of descendants that it has after some chosen time interval; they use

descendants rather than offspring in order to allow their model to go beyond one generation (and

we have already seen reasons for doing so). However, each organism is not assigned one

particular value; rather, fitness is treated as having a distribution of possible values, each with a

probability assigned to it. In each generation, of course, an organism will have a determinate

number of descendants. Again, I refer the reader to the authors' papers for the technical details,

but the basic idea is this. If we were to run a computer simulation of a population of organisms,

each with a probability distribution of possible offspring values, we could have the computer

assign each organism an actual number of offspring in each generation based on the organism's

probability distribution.[21] Offspring would be assigned a phenotype (and thus a distribution of

their possible offspring values) based on a distribution of possible offspring phenotypes assigned

to parents. With many runs of the simulation, we can thus see how different distributions

compare, i.e., whether *A*, with its entire distribution, is in fact fitter than *B* by whether it produces

a greater number of descendants in most of the simulations, for as many generations as we care

---

[20] Thanks to John Beatty for the suggestion to look at Rice's work.
[21] Rice and colleagues also describe an analytical solution, but I think the computer model (they have performed Monte Carlo simulations) is a little easier to understand.

to examine.

I think the approach of Rice and colleagues is extremely promising; in addition to potentially addressing the worries raised by Beatty and Finsen, it has at least three other benefits. First, they are offering not just a way to understand fitness values; they are also offering a new way to model selection processes. More traditional approaches, such as those based on the Hardy-Weinberg equation or the Price equation, tend to treat fitness as a fixed parameter. Thus, there has always been a bit of a disconnect between the way we have understood the concept of fitness (as probabilistic) and the way that the models have treated fitness (as fixed and determinate). The model that Rice and colleagues propose explicitly incorporates a probabilistic notion of fitness, thus making our concept and our model more consistent with one another and yielding a model that is more realistic as well. Again, as Darwin himself noted, the fitter organism may not in fact be the one that is the more successful. A second benefit of their way of treating fitness is that, by considering the full distribution instead of a limited set of moments, new results have been found. For example, Rice (2008) characterizes new situations where the expected direction of evolution is toward the phenotype with *higher* variance in fitness. A third benefit is that situations where fitnesses are fluctuating due to fluctuating environment, or where the actions of organisms are affecting their fitnesses (niche construction), are more easily accommodated.

Perhaps most importantly, however, the approach of Rice and colleagues vindicates Brandon's (1990) suggestion that we should separate what fitness *is* from how we assign *values* to it. Expected reproductive success is not the propensity interpretation of fitness and it never was; it was just one way of trying to grapple with probability distributions and assign fitness values.

Whether the particular approach of Rice and colleagues succeeds, or whether some variant is needed, shows that it is a *mathematical* problem as to how one compares probability distributions. Clearly, this is a difficult thing to do, but these difficult puzzles don't change the fact that it is the probability distributions which represent the value of an organism's fitness.

Fitness is an organism's propensity to survive and reproduce (based on its heritable physical traits) in a particular environment and a particular population over a specified number of generations.[22] That is what fitness *is*. A probability distribution of possible offspring contributions for particular organisms can be used to compare the fitness of organisms and to make predictions of future frequencies in the population. That is how fitness can be assigned a *value*.[23] As for the best way to compare probability distributions, I leave that to mathematicians and mathematical biologists.

5.4 Propensities are Causal

Propensities are usually understood to be causal; indeed, that they are causal is one of the features of propensities that philosophers of probability find challenging, because their causal nature implies that they must *not* obey the usual (Kolmogorov) probability calculus (Hájek 2012).[24] That is, propensities are usually held to be the cause of observed, actual frequencies; the propensity of a fair coin to land heads 50% of the time is the cause of the actual sequence of

---

[22] Fitness is also arguably a property of other biological entities, such as genes, groups, and species; however, here I will set aside debates over which entities are proper units of selection and in what sense.
[23] This solution, if sound, might make Sober's worries about incorporation of population size into our calculations moot. However, the reasons for considering the fitness of an organism to be relative to its population (as well as to its environment) still hold.
[24] A worry which I will not be addressing in this essay. The reader should note, however, that alternative probability calculuses have been developed.

heads and tails (e.g., HTTHHHTHHT). Yet Sober – in the same book where he endorsed the propensity interpretation – famously argued that fitness is "causally inert" (unlike some propensities, which he thinks are causal) even though natural selection is one of the causes of evolution (Sober 1984; see also Hodge 1987). On the other hand, Brandon and Beatty state that "[a]ccording to the propensity interpretation, the connection between the ability and the actual manifestation of the ability is a causal connection" (1984: 344). Who is right?

Sober denies that "overall fitness is causally efficacious" (1984: 89). The word "overall" is key here. The essence of his argument is this: the overall fitness of an organism is made up of many different abilities (in relation to its environment), such as the ability to avoid disease and the ability to escape predation. However, Sober argues, in a given survival event, both of those abilities may not be relevant:

> When an individual survives because it manages to avoid being caught by a predator, I can see how its *invulnerability to predation* was a causal factor contributing to its survival. Likewise, when an individual is infected by a disease but survives because of its immunity, I can see how its *invulnerability to disease* helped keep it alive. But an organism's overall fitness–its high probability of surviving, no matter what cause of mortality may present itself–strikes me as being causally inert (Sober 1984: 91).

The overall fitness of an organism is, Sober asserts, a disjunction of abilities. One of them may be causally efficacious in a particular case, but the disjunction is not. If an organism escapes a

predator, its immunity to disease played no causal role,[25] and so, Sober suggests, it does not make sense to say that its overall fitness played a causal role.

However, it's not clear that "overall fitness" is the concept of fitness at stake; rather, it is arguably *trait fitness* that we ought to be considering.[26] Mills and Beatty state that biologists are generally talking about the relative fitness of types (genotypes or phenotypes), where organisms share a gene or a trait but vary in other respects; what biologists are trying to do is to "explain the evolution and/or persistence of a gene or its phenotypic manifestation in a temporally extended population" by showing "that possessors of the gene or trait were generally better able to survive and reproduce than possessors of alternate traits or genes" (1979: 346). Indeed, many of the classic examples of natural selection focus not on an organism's fitness, but on the fitness of organisms given a certain specified trait: the beak of a finch, the color of a moth, the banding pattern of a snail. The other traits are backgrounded and assumed not to be correlated with the trait in question. To elaborate further, suppose that (as a defeasible hypothesis), a population had some organisms with trait *X* and other organisms with a different trait *Y*, but that either 1) other traits were distributed randomly across those two subgroupings, or 2) there was no difference in the distributions of other traits across the two subgroupings. These situations would be analogous to a randomized experiment or an experiment with matching, respectively. Under these conditions, trait fitness can be shown to be a causal factor contributing to organisms' survival in the population.

Note the focus on *relative* fitness (in the comparative sense) in the explanatory practices of

---

[25] Note that to say that the traits of an organism can be separated causally is a substantive empirical assumption, given the high degree of integration that organisms have.

[26] See also Ramsey (2013) for a discussion of trait fitnesses in the context of Sober's arguments.

biologists. Indeed, it *is* a mistake to say that the fitness of individual organisms and traits (*fitness₁* and *fitness₂,* respectively) play a causal role *in natural selection*. Only relative fitness (*relative fitness₂*) matters for natural selection. The ability of a type to produce 100 offspring will be selectively favored only if other types in the population have a lesser ability; if others have a greater ability, the first type will be selected against, and if there are no fitness differences at all, there will be no natural selection at all.

If this explanation of how trait fitness can be a causal factor in the population is right, then Sober's argument is simply not applicable to the propensity interpretation of fitness, which is not overall fitness at all but rather the fitness of organisms with a certain trait, which Sober seems to grant is unproblematically causal[27] (again, recall that to claim that there is a propensity is just to claim that there is an underlying physical basis for the behavior). That is, the propensity of fitness would be causally efficacious, not causally inert.

Furthermore, even if it turned out that the overall fitness of an *organism* is causally inert, it might still be the case that overall fitness *differences* between organisms are causally efficacious. One particular organism may survive a predator rather than survive a disease (supposing that the organism was never exposed to the disease in question), but in the population as a whole, one would expect some organisms to survive predators, others to survive disease, etc. Thus, the overall fitness differences – more properly a property of the population rather than a property of any individual organisms – would be causally relevant to the changes in the distribution of

---

[27] But see Sober (2013).

organisms over time.[28] It still might turn out that in a particular generation, no organism was

exposed to a given disease – but here it is important to keep in mind that since fitness is relative

to the environment, an environment that lacks the relevant disease is one in which immunity does

not contribute to overall fitness. And an environment with very little risk of disease is one in

which immunity makes very little contribution to overall fitness.

6. Conclusions

I have argued that the propensity interpretation of fitness, properly understood, not only solves

the explanatory circularity problem and the mismatch problem, but can also withstand the

Pandora's box full of problems that have been thrown at it.[29] Fitness is the propensity (i.e.,

probabilistic ability, based on heritable physical traits) for organisms or types of organisms to

survive and reproduce in particular environments and in particular populations for a specified

number of generations; if greater than one generation, "reproduction" includes descendants of

descendants. Fitness values can be described in terms of distributions of propensities to produce

varying number of offspring and can be modeled for any number of generations using computer

simulations, thus providing both predictive power and a means for comparing the fitness of

different phenotypes. Fitness is a causal concept, most notably at the population level, where

fitness differences are causally responsible for differences in reproductive success. Relative

fitness is ultimately what matters for natural selection.

---

[28] This general point is argued for in more detail in Millstein (2006), but without referring to fitness directly; Sober (2013) seems to endorse it.

[29] See Drouet and Merlin (forthcoming) for a recent argument that fitness should not be understood as propensities, and Pence and Ramsey (2013) for an alternative defense of it.

More generally, the above examination of fitness in the theory of natural selection implies the following lessons for understanding probability in other parts of evolutionary biology and other biological sciences.

First, if fitness and thus natural selection are probabilistic, then given the ubiquity of natural selection as an evolutionary process (not to mention random drift and other probabilistic evolutionary processes), we should *expect* to find probabilities throughout biology. Philosophers have only begun to scratch the surface in seeking out probabilities in the various areas of biology discussed at the outset of this paper. This means that there are potentially many analogous discussions to the one presented here; these should be explored. Furthermore, any failure to find the expected probabilities raises questions. If the study of an area of biology does not seem to invoke probabilities, why not? Does the aspect of organisms under study truly behave in a determinate way? Or are there other (pragmatic?) reasons why probabilities have not been invoked?

Second is a recommendation for the general approach that Mills and Beatty (1979) took, which has proved to be so fruitful. The stated definitions of biologists induced an apparent puzzle, the explanatory circularity problem. It is also often the case that the stated definitions of biologists can conflict. Thus, a promising way to understand concepts is to look at the way that biologists *use* their terms – that is, to look at actual biological practice. This holds for any sort of concept, but might be particularly useful for probabilistic ones, which can be difficult to express.

Third is not to lose sight of the original problems that motivate the understanding of a

probabilistic concept. Technical issues (such as the mathematical ones discussed in this paper) can be intriguing for philosophers and are often productive, but they can also lead away from the core issues and cause one to misunderstand where the problem really lies. Philosophy of biology is most productive when it keeps both philosophy and actual biological practice in the forefront.

Finally, I hope to have shown that thinking about interpretations of probability in general, as well as the particular details of the invoked interpretation of probability, can be useful. Interpretations of probability provide criteria of things to look for and questions to ask. Figuring out which interpretations of probability are appropriate for different areas of biology is thus an essential first step. In the case of the propensity interpretation, it is important to understand that propensities are relational, that they are causal, and that they are always for a specified behavior; addressing these criteria leads to a better understanding of the nature of fitness. Other criteria can be identified for other interpretations of probability, such as whether degrees of belief (such as the uncertainties in medicine) are an inherent part of the invoked probabilities. In other words, interpretations of probability are more than just the answers to interesting philosophical questions; they can be useful tools.

selection; and of course to John Beatty for always gently pushing me to think harder about all things related to evolution and chance.

References

Abrams, Marshall. 2006. Infinite Populations and Counterfactual Frequencies in Evolutionary Theory. *Studies in History and Philosophy of Biological and Biomedical Sciences* 37:256-268.

---. 2007. Fitness and Propensity's Annulment. *Biology and Philosophy* 22:115-130.

Ariew, André, and Zachary Ernst. 2009. What Fitness Can't Be. *Erkenntnis* 71:289-301.

Ariew, André, and Richard C. Lewontin. 2004. The Confusions of Fitness. *British Journal for the Philosophy of Science* 55:347-363.

Beatty, John. 1984. Chance and Natural Selection. *Philosophy of Science* 51:183-211.

Beatty, John, and Susan Finsen. 1989. Rethinking the Propensity Interpretation: A Peek Inside Pandora's Box. In *What the Philosophy of Biology Is*, edited by M. Ruse. Dordrecht: Kluwer.

Bouchard, Frédéric. 2011. Darwinism without populations: a more inclusive understanding of the 'Survival of the Fittest'. *Studies in History and Philosophy of Biological and Biomedical Sciences* 42:106-114.

Bouchard, Frédéric, and Alex Rosenberg. 2004. Fitness, Probability, and the Principles of Natural Selection. *British Journal for the Philosophy of Science* 55:693-712.

Brandon, Robert, and John Beatty. 1984. The Propensity Interpretation of 'Fitness'–No Interpretation Is No Substitute. *Philosophy of Science* 51 (2):342-347.

Brandon, Robert N. 1978. Adaptation and Evolutionary Theory. *Studies in the History and Philosophy of Science* 9:181-206.

---. 1990. *Adaptation and Environment*. Princeton: Princeton University Press.

Burian, Richard M. 1983. Adaptation. In *Dimensions of Darwinism*, edited by M. Greene. Cambridge: Cambridge University Press.

Colyvan, Mark. 2005. Probability and Ecological Complexity. *Biology and Philosophy* 20:869-879.

Craver, Carl. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Clarendon Press.

Darwin, Charles. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 1st ed. London: John Murray.

---. [1871] 1981. *The Descent of Man, and Selection in Relation to Sex*. First ed. Princeton: Princeton University Press.

Drouet, Isabelle, and Francesca Merlin. Forthcoming. The Propensity Interpretation of Fitness and the Propensity Interpretation of Probability. *Erkenntnis*.

Djulbegovic, Benjamin, Iztok Hozo, and Sander Greenland. 2011. Uncertainty in Clinical Medicine. In *Philosophy of Medicine*, ed. Fred Gifford. Oxford: Elsevier.

Fox Keller, Evelyn. 1987. Reproduction and the Central Project of Evolutionary Theory. *Biology and Philosophy* 2:383-396.

Gayon, Jean. 1998. *Darwinism's Struggle for Survival: Heredity and the Hypothesis of Natural Selection*. Cambridge: Cambridge University Press.

Gillespie, John H. 1974. Natural Selection for Within-Generation Variance in Offspring Number. *Genetics* 76:601-606.

---. 1977. Natural Selection for Variances in Offspring Numbers–A New Evolutionary Principle. *American Naturalist* 111:1010-1014.

Gillies, Donald. 2000. *Philosophical Theories of Probability*: Routledge.

Hájek, Alan. 2012. Interpretations of Probability. *The Stanford Encyclopedia of Philosophy* Summer 2012 Edition.

Hodge, M. J. S. 1987. Natural Selection as a Causal, Empirical, and Probabilistic Theory. In *The Probabilistic Revolution*. Cambridge: MIT Press.

Jackson, Jr., John P. manuscript. The Survival of the Unfit: Was Eugenics a Darwinian Discourse?

Kitcher, Philip. 1978. Why Not the Best? In *The Latest on the Best: Essays on Evolution and Optimality*, edited by J. Dupré. Cambridge: MIT Press.

Lennox, James G., and Bradley E. Wilson. 1994. Natural Selection and the Struggle for Existence. *Studies in the History and Philosophy of Science* 25 (1):65-80.

Matthen, Mohan, and André Ariew. 2002. Two Ways of Thinking About Fitness and Natural Selection. *The Journal of Philosophy* XCIX (2):55-83.

Mayr, Ernst. 1988. *Toward A New Philosophy of Biology. Observations of an Evolutionist"*. Cambridge: Harvard University Press.

Merlin, Francesca. 2010. Evolutionary Chance Mutation: A Defense of the Modern Synthesis' Consensus View. *Philosophy and Theory in Biology* 2:e103:1-22.

Mills, Susan K., and John H. Beatty. 1979. The Propensity Interpretation of Fitness. *Philosophy of Science* 46 (2):263-286.

Millstein, Roberta L. 2002. Are Random Drift and Natural Selection Conceptually Distinct? *Biology and Philosophy* 17 (1):33-53.

---. 2003a. Interpretations of Probability in Evolutionary Theory. *Philosophy of Science* 70 (5):1317-1328.

---. 2003b. How Not to Argue for the Indeterminism of Evolution: A Look at Two Recent Attempts to Settle the Issue. In *Determinism in Physics and Biology*, edited by A. Hüttemann. Paderborn: Mentis.

---. 2006. Natural Selection as a Population-Level Causal Process. *The British Journal for the Philosophy of Science* 57 (4):627-653.

---. 2009. Populations as Individuals. *Biological Theory* 4 (3):267-273.

---. 2010. The Concepts of Population and Metapopulation in Evolutionary Biology and Ecology. In *Evolution Since Darwin: The First 150 Years*, edited by M. A. Bell, D. J. Futuyma, W. F. Eanes and J. S. Levinton. Sunderland, MA: Sinauer.

---. 2011. Chances and Causes in Evolutionary Biology: How Many Chances Become One Chance. In *Causality in the Sciences*, eds. P. McKay Illari, F. Russo, and J. Williamson, 425-444. Oxford: Oxford University Press.

---. 2012. Darwin's Explanation of Races by Means of Sexual Selection. *Studies in History and Philosophy of Biological and Biomedical Sciences* 43:627-633.

Millstein, Roberta L., and Robert A. Skipper. 2007. Population Genetics. In *The Cambridge Companion to the Philosophy of Biology*, edited by D. Hull and M. Ruse. Cambridge: Cambridge University Press.

Pence, Charles H., and Grant Ramsey. 2013. A New Foundation for the Propensity Interpretation of Fitness. *British Journal for the Philosophy of Science* 64:851-881.

Pigliucci, Massimo, and Jonathan Kaplan. 2006. *Making Sense of Evolution: The Conceptual Foundations of Evolutionary Biology* Chicago: University of Chicago Press.

Popper, Karl R. 1959. The Propensity Interpretation of Probability. *The British Journal for the Philosophy of Science* 10 (37):25-42.

Ramsey, Grant. Can fitness differences be a cause of evolution? *Philosophy and Theory in Biology* 5:e401.

Rice, Sean. 2008. A stochastic version of the Price equation reveals the interplay of deterministic and stochastic processes in evolution. *BMC Evolutionary Biology* 8:262-.

Rice, Sean, and Anthony Papadopoulos. 2009. Evolution with Stochastic Fitness and Stochastic Migration. *PLoS ONE* 4 (10):1-12.

Rice, Sean, Anthony Papadopoulos, and John Harting. 2011. Stochastic Processes Driving Directional Evolution. In *Evolutionary Biology–Concepts, Methods, Macroevolution*, edited by P. Pontarotti. Berlin: Springer-Verlag.

Richardson, Robert C., and Richard M. Burian. 1992. A Defense of Propensity Interpretations of Fitness. *PSA 1992: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1:349-362.

Rosenberg, Alex. 1985. *The Structure of Biological Science*. Cambridge: Cambridge University Press.

Rosenberg, Alex, and Mary B. Williams. 1986. Discussion: Fitness as a Primitive and Propensity. *Philosophy of Science* 53:412-418.

Scriven, M. 1959. Explanation and Prediction in Evolutionary Theory. *Science* 130 (477-482).

Sober, Elliott. 1984. *The Nature of Selection*. Chicago: University of Chicago Press.

---. 2001. The Two Faces of Fitness. In *Thinking About Evolution: Historical, Philosophical, and Politcal Perspectives*, edited by R. S. Singh, C. B. Krimbas, D. B. Paul and J. Beatty. Cambridge: Cambridge University Press.

---. 2013. Trait fitness is not a propensity, but fitness variation is. *Studies in History and Philosophy of Biological and Biomedical Sciences*. 44:336-341.

Walsh, Denis M. 2007. The Pomp of Superfluous Causes: The Interpretation of Evolutionary Theory. *Philosophy of Science* 74:281-303.

Waters, Kenneth C. 1986. Natural Selection Without Survival of the Fittest. *Biology and Philosophy* 1:207-225.